Segmenting the Periocular Region using a Hierarchical Graphical Model Fed by Texture / Shape Information and Geometrical Constraints*

Hugo Proença, João C. Neves and Gil Santos
IT - Instituto de Telecomunicações
University of Beira Interior, Portugal
{hugomcp,jcneves, gsantos}@di.ubi.pt

Abstract

Using the periocular region for biometric recognition is an interesting possibility: this area of the human body is highly discriminative among subjects and relatively stable in appearance. In this paper, the main idea is that improved solutions for defining the periocular region-of-interest and better pose / gaze estimates can be obtained by segmenting (labelling) all the components in the periocular vicinity. Accordingly, we describe an integrated algorithm for labelling the periocular region, that uses a unique model to discriminate between seven components in a single-shot*: iris, sclera, eyelashes, eyebrows, hair, skin and glasses. Our solution fuses texture / shape descriptors and geometrical constraints to feed a two-layered graphical model (Markov Random Field), which energy minimization provides a robust solution against uncontrolled lighting conditions and variations in subjects pose and gaze.*

1. Introduction

Motivated by the pioneering work of Park *et* al. [14], the concept of periocular recognition has been gaining relevance in the biometrics literature, particularly for uncontrolled data acquisition setups. For such cases, the idea is that - apart the iris - additional discriminating information can be obtained from the skin and sclera textures, and the shape of eyelids, eyelashes and eyebrows.

Most of the relevant periocular recognition algorithms work in a *holistic* way, i.e., they define a region-of-interest (ROI) around the eye and apply a feature encoding strategy independently of the biological component at each position. The exceptions (e.g., [17] and [6]) regard the iris and the sclera components, for which specific feature encoding / matching algorithms are used. This observation leads that some components (e.g., hair or glasses) might be erroneously taken into account and bias the recognition process.

The automatic labelling (segmentation) of the components in the periocular region has - at least - two obvious advantages: it enables to define better ROIs and conducts to more accurate estimates of subjects' pose and gaze. Hence, this paper describes an image labelling algorithm for the periocular region that discriminates between seven components (iris, sclera, eyelashes, eyebrows, hair, skin and glasses), according to a model composed of two phases:

1. seven non-linear classifiers running at the pixel level are inferred from a training set, and provide the posterior probabilities for each image position and class of interest. Each classifier (neural network) is specialized in detecting one component and receives local statistics (texture and shape descriptors) from the input data;

2. the posteriors based on data local *appearance* are combined with geometric constraints and components' adjacency priors, to feed a hierarchical Markov Random Field (MRF), composed of a *pixel* and a *component* layer. MRFs are a classical tool for various computer vision problems, from image segmentation (e.g., [10]), image registration (e.g., [8]) to object recognition (e.g., [5]). Among other advantages, they provide non-causal models with isotropic behavior and faithfully model a broad range of local dependencies. The model proposed in this paper inherits some insights from previous works that used shape priors to constraint the final model (e.g., [3]) and multiple layered MRFs (e.g., [19]).

To illustrate the usefulness of the proposed algorithm, we compare the effectiveness of the Park *et* al.'s [14] recognition method, when using the ROI as originally described and according to an improved version, that considers the center of mass of the cornea as reference point (less sensitive to gaze) and avoids that hair and glasses inside the ROI are considered in feature encoding / matching. The

*This work was supported by FCT project PEst-OE/EEI/LA0008/2013

observed improvements in performance anticipate other benefits that can be attained by labelling the periocular region before recognition: pose / gaze estimates based in the labelled data and development of component-specific feature encoding / matching strategies.

The remainder of this paper is organized as follows: Section 2 summarizes the most relevant periocular recognition algorithms. Section 3 provides a description of the proposed model. Section 4 regards the empirical evaluation and the corresponding results. Finally, the conclusions are given in Section 5.

2. Periocular Recognition: Literature Review

The first work in this field was published in 2009, due to Park *et* al. [14]. They characterised the periocular region by local binary patterns (LBP), histograms of oriented gradients (HOG) and scale-invariant feature transforms (SIFT), fused at the score level. Subsequently, the same authors [13] described additional factors that affect performance, including segmentation inaccuracies, partial occlusions and pose. Woodard *et* al. [20] observed that fusing the responses from periocular and iris recognition modules improves performance with respect to each system considered individually. Bharadwaj *et* al. [4] fused a global descriptor based on five perceptual dimensions (image naturalness, openness, roughness, expansion and ruggedness) to circular LBPs. The Chi-square distances from both types of features were finally fused at the score level. Ross *et* al. [16] handled challenging deformed samples, using probabilistic deformation models and maximum-a-posteriori estimation filters. Also concerned about robustness, Woodard *et* al. [21] represented the skin texture and color using separate features, that were fused in the final stage of the processing chain. Tan *et* al. [18] proposed a method that got the best performance in the *NICE: Noisy Iris Challenge Evaluation*[1]. contest. This method is actually a periocular recognition algorithm: texton histograms and semantic rules encode information from the surroundings of the eye, while ordinal measures and color histograms encode the iris data. Oh *et* al. [9] combined sclera and periocular features: directional periocular features were extracted by structured random projections, complemented by a binary representation of the sclera. Tan and Kumar [17] fused iris information (encoded by Log-Gabor filters) to an over-complete representation of the periocular region (LBP, GIST, HOG and Leung-Malik Filters). Both representations were matched independently and fused at the score level.

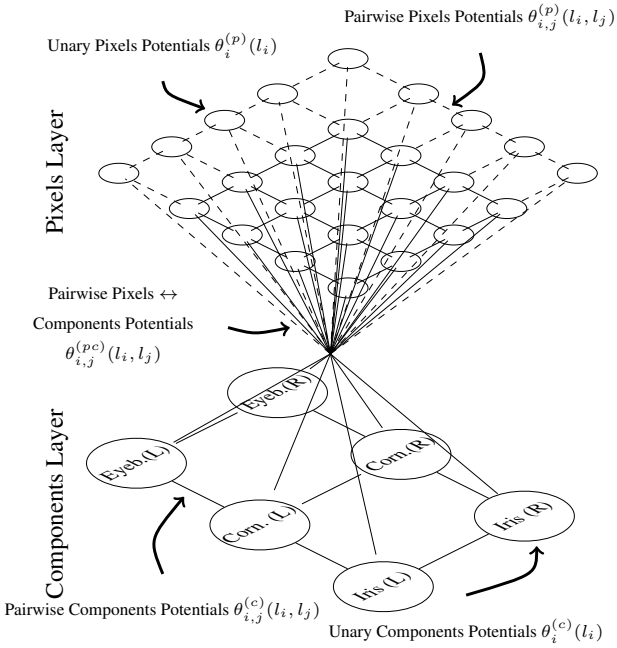

Figure 1. Structure of the MRF that segments the periocular region.

3. Proposed Method

As Fig. 1 illustrates, the proposed MRF is composed of two layers: one works at the *pixel* level, with a bijection between each image pixel and a vertex in the MRF. The second layer regards the major *components* in the periocular vicinity, with six vertices representing the eyebrows, irises and corneas from both sides of the face. The insight behind this structure is that the pixels layer mainly regards the data appearance, while the components layer represents the geometrical constraints in the problem and assures that the generated solutions are biologically plausible.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph representing a MRF, composed of a set of t_v vertices $\mathcal{V}$, linked by t_e edges $\mathcal{E}$. Let t_p be the number of vertices in the *pixels* layer and let t_c be the number of vertices in the *components* layer, such that $t_v = t_p + t_c$. Let $\mathfrak{C}(x, y)$ denote the biological component at position (x, y) of an image and $\mathfrak{T}_j$ be the component's *type* of the j^{th} component node: either 'iris', 'cornea' or 'eyebrow'.

The MRF is a representation of a discrete latent random variable $\boldsymbol{L} = \{L_i\}, \forall i \in \mathcal{V}$, where each element L_i takes one value l_i from a set of labels. Let $\boldsymbol{l} = \{l_1, \ldots, l_{t_p}, l_{t_p+1}, \ldots, l_{t_p+t_c}\}$ be one configuration of the MRF. In our model, every component node is directly con-

[1] http://nice2.di.ubi.pt/

nected to each pixel node and the pixel nodes are connected to their horizontal / vertical neighbors (4-connections). Also, the edges between component nodes correspond to geometrical / biological constraints in the periocular region: the nodes representing both irises, corneas and eyebrows are connected, as do the iris, cornea and eyebrow nodes of the same side of the face. Note that the proposed model does not use high-order potentials. Even though there is a point in Fig. 1 that joins multiple edges, it actually represents overlapped pairwise connections between one component and one pixel vertex.

The energy of a configuration $\boldsymbol{l}$ of the MRF is the sum of the unary $\theta_i(l_i)$ and pairwise $\theta_{i,j}(l_i, l_j)$ potentials:

$$E(\boldsymbol{l}) = \sum_{i \in \mathcal{V}} \theta_i(l_i) + \sum_{(i,j) \in \mathcal{E}} \theta_{i,j}(l_i, l_j). \quad (1)$$

According to this formulation, labelling an image is equivalent to infer the random variables in the MRF by minimizing its energy:

$$\hat{\boldsymbol{l}} = \arg\min_{\boldsymbol{l}} E(\boldsymbol{l}), \quad (2)$$

where $\{\hat{l}_1, \ldots, \hat{l}_{t_p}\}$ are the labels of the pixels and $\{\hat{l}_{t_p+1}, \ldots, \hat{l}_{t_p+t_c}\}$ specify the components' parameterizations. In this paper, the MRF was optimized according to the Loopy Belief Propagation [7] algorithm. Even though it is not guaranteed to converge to global minimums on loopy non-submodular graphs (such as our MRF), we concluded that the algorithm provides visually pleasant solutions most of the times. As future work, we plan to evaluate the effectiveness of our model according to more sophisticated energy minimization algorithms (e.g., sequential tree-reweighed message passing [11]).

3.1. Feature Extraction

Previous works reported that the hue and saturation channels of the HSV color space are particularly powerful to detect the sclera [15], whereas the red / blue chroma values provide good separability between the skin and non-skin pixels [1]. Also, the iris color triplets are typically distant from the remaining periocular components and there is a higher amount of information in patches of the eyebrows and hair regions than in the remaining components. Accordingly, a feature set at the pixel level is extracted, composed of 34 elements (Fig. 2): {red, green and blue channels (RGB); hue, saturation and value channels (HSV); red and blue chroma (yCbCr); LBP and entropy in the value channel}, all averaged in square patches of side $\{3, 5, 7\}$ around the central pixel. Also, the convolution between the value channel and a set of Gabor kernels $\boldsymbol{G}$ complements the feature set:

$$\boldsymbol{G}[x, y, \omega, \varphi, \sigma] = \exp\left[\frac{-x^2 - y^2}{\sigma^2}\right] \exp[2\pi\omega i \Phi] \quad (3)$$

being $\Phi = x\cos(\varphi) + y\sin(\varphi)$, ω the spatial frequency, φ the orientation and σ the standard deviation of an isotropic Gaussian kernel ($\omega \in \{\frac{3}{2}, \frac{5}{2}\}, \varphi \in \{0, \frac{\pi}{2}\}, \sigma = 0.65\omega$).

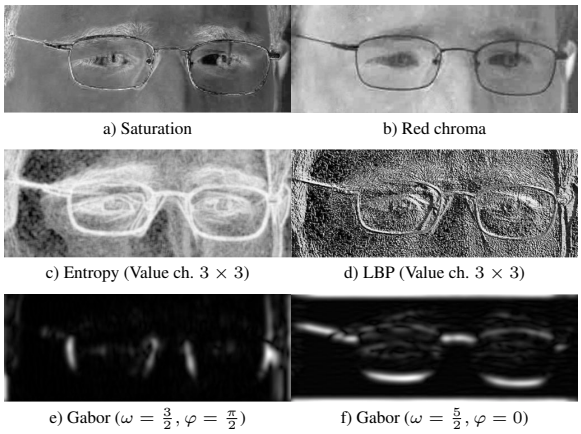

Figure 2. Illustration of the discriminating power of the features extracted, for the seven classes considered in this paper.

3.2. Unary Potentials

Let $\gamma : \mathbb{N}^2 \rightarrow \mathbb{R}^{34}$ be the feature extraction function, that for each image pixel (x, y) returns a feature vector $\gamma(x, y) \in \mathbb{R}^{34}$. Let $\Gamma = [\gamma(x_1, y_1), \ldots, \gamma(x_n, y_n)]^T$ be a $n \times 34$ matrix extracted from a training set, that is used to learn seven non-linear binary classification models, each one specialized in detecting a component (class) $\omega_i \in$ {Iris, Sclera, Eyebrows, Eyelashes, Hair, Skin, Glasses}. Let $\eta_i : \mathbb{R}^{34} \rightarrow [0, 1]$ be the response of the ith non-linear model, used to obtain the likelihood of class ω_i: $p\Big(\eta_i\big(\gamma(x, y)\big)|\omega_i\Big)$. According to the Bayes rule, assuming equal priors, the posterior probability functions are given by:

$$P\Big(\omega_i|\eta_i\big(\gamma(x, y)\big)\Big) = \frac{P\Big(\eta_i\big(\gamma(x, y)\big)|\omega_i\Big)}{\sum_{j=1}^{7} P\Big(\eta_j\big(\gamma(x, y)\big)|\omega_j\Big)}. \quad (4)$$

The unary potentials of each vertex in the pixels layer are defined as $\theta_i^{(p)}(l_i) = 1 - p\Big(\omega_i|\eta_i\big(\gamma(x, y)\big)\Big)$.

Each label in the components layer represents a parameterisation of an ellipse (found by the Random Elliptical Hough Transform (REHT)) [2] that roughly models the eyebrows, corneal or iris regions. Starting from images labelled by the index of the maximum posterior probability

$I_m(x,y) = \arg\max_j p\Big(\omega_j | \eta_j\big(\gamma(x,y)\big)\Big)$ (upper image in Fig. 3), a binary version per component can be obtained (bottom images in Fig. 3):

$$I_{m_i}(x,y) = \begin{cases} 1 & \text{, if } I_m(x,y) = i \\ 0 & \text{, otherwise} \end{cases} \quad (5)$$

The output of the REHT algorithm in $I_{m_i}(x,y)$ gives the unary potential of the component vertices: $\theta_i^{(c)}(l_i) = -\log\big(\kappa(i)\big), \forall i \in t_{p+1}, \ldots, t_{p+c}$, being $\kappa(i)$ the votes returned by the REHT for the ith ellipse parameterisation.

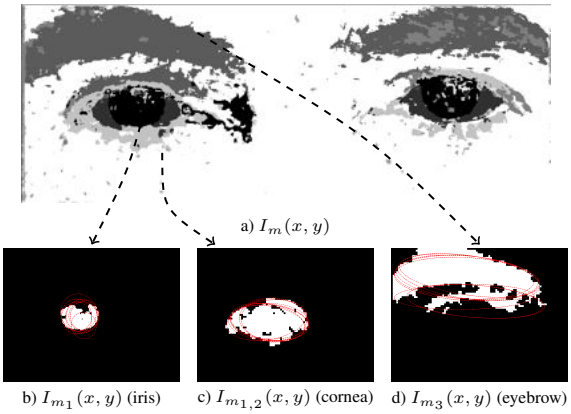

Figure 3. (Upper row) Example of an image labelled by the maximum of the posteriors given by the classification models $\eta_i\big(\gamma(x,y)\big)$. The red ellipses in the bottom images represent the parameterisations returned by the REHT algorithm for the left iris, cornea and eyebrow.

3.3. Pairwise Potentials

There are three types of pairwise potentials in our model: 1) between two pixel nodes; 2) between two component nodes; and 3) between a pixel and a component. The pairwise potential between pixel nodes spatially adjacent $\theta_{i,j}^{(p)}(l_i, l_j)$ is defined as the prior probability of observing labels l_i, l_j in adjacent positions of a training set (e.g., it is much more probable that an "eyebrow" pixel is adjacent to a "skin" pixel than to an "iris" one):

$$\theta_{i,j}^{(p)}(l_i, l_j) = \frac{1}{\alpha_0 + P(\mathfrak{C}(x',y') = \omega_i, \mathfrak{C}(x,y) = \omega_j)}, \quad (6)$$

where $P(.,.)$ is the joint probability, (x', y') and (x, y) are 4adjacent positions and $\alpha_0 \in \mathbb{R}^+$ avoids infinite costs (likewise, all α_i terms below are regularization terms).

The pairwise potentials between component nodes consider the geometrical constraints in the periocular area, i.e., enforce that the irises are inside the cornea, and below the eyebrows. Also, both irises, corneas and eyebrows should have similar vertical coordinate and similar size. Let $(x_i, y_i, a_i, b_i, \varphi_i)$ be the ith parameterisation of an ellipse, being (x_i, y_i) the ellipse centre, (a_i, b_i) its major / minor axes and φ_i the rotation. For pairs of nodes of the same type ($\mathfrak{T}_i = \mathfrak{T}_j$), similar vertical coordinates and similar sizes are privileged:

$$\theta_{i,j}^{(c1)}(l_i, l_j) = \alpha_1 |y_i - y_j| + \alpha_2 |a_i + b_i - a_j - b_j|. \quad (8)$$

For edges connecting the cornea (ith node) and the eyebrow (jth node) we privilege similar horizontal coordinates and locations having the eyebrow above the cornea:

$$\theta_{i,j}^{(c2)}(l_i, l_j) = \alpha_3 |x_i - x_j| + \alpha_4 \max\{0, y_i - y_j\}. \quad (9)$$

Regarding the iris / cornea pairwise potentials, we penalize parameterizations with portions of the iris outside the cornea:

$$\theta_{i,j}^{(c3)}(l_i, l_j) = \alpha_5\Big(1 - \frac{\sum_{x_i}\sum_{y_i} \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j)}{\sum_{x_i}\sum_{y_i} 1}\Big), \quad (10)$$

being (x_i, y_i) a pixel labelled as iris and $\psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j)$ an indicator function that verifies if that position is inside the ellipse defined by the jth parameterisation (7). Overall, the pairwise potentials in the components layer are defined as:

$$\theta_{i,j}^{(c)}(l_i, l_j) = \sum_{k=1}^{3} \theta_{i,j}^{(ck)}(l_i, l_j). \quad (11)$$

Lastly, the pairwise potentials between pixels and components enforce that pixels inside a component parameterisation are predominantly labelled by the value that corresponds to that type of node, whereas pixels outside that parameterisation should have label different of the component's type. Let (x_{jk}, y_{jk}) be the coordinates of the ellipse defined by the jth parameterization. The pairwise cost between the ith pixel node and the jth component node is given by:

$$\theta_{i,j}^{(pc)}(l_i, l_j) = \begin{cases} \min_k ||(x_i, y_i) - (x_{jk}, y_{jk})||_2, \text{if } l_i \in \mathfrak{T}_j \\ \quad \text{and } \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j) = 0 \\ 0, \text{if } l_i \notin \mathfrak{T}_j \\ \quad \text{and } \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j) = 0 \\ 0, \text{if } l_i \in \mathfrak{T}_j \\ \quad \text{and } \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j) = 1 \\ \max_k ||(x_i, y_i) - (x_{jk}, y_{jk})||_2, \text{if } l_i \notin \mathfrak{T}_j \\ \quad \text{and } \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j) = 1 \end{cases}, \quad (12)$$

where $||.||$ is the Euclidean distance.

$$\psi(x,y,x_i,y_i,a_i,b_i,\varphi_i) = \begin{cases} 1 & \text{, if } \frac{\left(\cos(\varphi_i)(x-x_i)+\sin(\varphi_i)(y-y_i)\right)^2}{a_i^2} + \frac{\left(\sin(\varphi_i)(x-x_i)+\cos(\varphi_i)(y-y_i)\right)^2}{b_i^2} \leq 1 \\ 0 & \text{, otherwise} \end{cases} \quad (7)$$

4. Experiments

Our experiments were carried out in a data set composed of 5,551 visible-light images (with resolution 800 $\times$ 300) containing the periocular regions from both sides of the face. These images were the source for the UBIRIS.v2 dataset: they were collected in indoor unconstrained lighting environments and feature significant variations in scale, subjects' pose and gaze. For learning / evaluation purposes, 200 images were manually labelled, covering the seven classes we aim to deal with. This set was divided into two disjoint parts: 1) one used to learn the classification models and to estimate the prior unary / pairwise costs of the MRF; and 2) the complementary part served for quantitative performance evaluation.

To obtain the seven classification models, we used feed-forward neural networks with three layers and $\{34:17:1\}$ topology, with *tan-sigmoid* transfer functions in the input and hidden layers and linear transfer functions in the output layer. The learning sets were always balanced (random sampling) and the Resilient Back-propagation algorithm used to learn the classifiers. Regarding the MRF optimization, every image was resized to 200 $\times$ 75 pixels, i.e., $t_p = 15,000$ in our MRFs. Also, $\boldsymbol{\alpha} = \{0.01, 1, 2, 10, 10\}$.

4.1. Segmentation Performance

Fig. 4 illustrates the results typically attained by the proposed model. Their visual coherence is evident, where regions labelled as hair appear in pink, eyebrows in yellow, irises in green, eyelashes in black, sclera in blue and glasses in blueberry color. Also, solutions were biologically plausible in the large majority of the cases, for various hairstyles, and different subjects poses / gazes. A particularly interesting performance was observed for glasses, where the algorithm attained remarkable results for various types of frames. This was probably due to the fact that glasses were the unique non-biological component among the classes considered, which might had increased their dissimilarity with respect to the remaining components.

In opposition, the most concerning cases happened when the eyebrows and the hair were overlapped (bottom-right image in Fig. 4) . Also, for heavily deviated gazes, the sclera was sometimes under-segmented (typically, by non-detecting the less visible side). In opposition, eyelashes tended to be over-segmented, with isolated eyelashes being grouped in large eyelash regions, which might be due to excessive pairwise cost for observing different labels in adjacent positions of the pixels layer.

Labeling Error	NN (%)		MRF (%)	
Component	FP	FN	FP	FN
Iris	1.12 ± 0.29	9.06 ± 1.80	0.17 ± 0.03	2.61 ± 0.51
Sclera	1.61 ± 0.49	5.17 ± 0.83	0.19 ± 0.03	3.60 ± 0.82
Eyebrows	2.20 ± 0.40	6.93 ± 0.95	0.79 ± 0.28	2.25 ± 0.46
Eyelashes	1.47 ± 0.38	5.12 ± 1.13	0.93 ± 0.23	0.62 ± 0.53
Hair	3.16 ± 0.56	6.74 ± 1.27	1.26 ± 0.30	3.09 ± 0.88
Skin	4.10 ± 1.03	4.09 ± 0.69	2.63 ± 0.43	3.86 ± 1.01
Glasses	1.08 ± 0.22	5.03 ± 1.45	0.06 ± 0.01	0.60 ± 0.09

Table 1. Average pixel labelling errors per component, when considering exclusively the $\arg\max_j p\Big(\omega_j|\eta_j\big(\gamma(x,y)\big)\Big)$ value (NN column) and with the proposed MRF model (MRF column).

It should be noted that α_i were found in an empirical and independent way, i.e., no exhaustive evaluation of combined configurations was carried out, nor any parameter optimization algorithm was used, which also points for the robustness of the proposed model against sub-optimal parameterizations. Table 1 gives the error rates per class, when considering exclusively the first phase of our model (maximum of the posterior probabilities, column "NN") and the full processing chain (MRF optimization, column "MRF"). In this table, FP stands for the false positives rate, whereas FN refers to the false negatives rate. In all cases, it is evident that the MRF substantially lowered the labeling error rates, essentially by imposing smoother responses and constraining the range of biologically acceptable solutions.

As the machine learning algorithm described in this paper is *supervised*, it is important to perceive its variations in performance with respect to the amount of learning data used to create the classification models and the prior unary / pairwise potentials. To this end, performance was compared while varying the number of images used in learning, and keeping constant the number of images used in performance evaluation (to assure comparable bias / variance scores). Figure 5 expresses the results: the horizontal axis gives the number of learning images used and the vertical axis is the corresponding pixel classification error, with the corresponding 95% confidence intervals. We observed that when more than 35 images were used in learning, the pixel classification errors tend to converge. This is evident in terms of the absolute error values and of the narrowness of the confidence intervals.

Figure 4. Examples of the segmented periocular regions. "Hair" class is represented by the pink color, "Eyebrows" appear in yellow, "Iris" in green, "Sclera" in blue, "Glasses" in blueberry and "Eyelashes" in gray. Pixels classified as "Skin" are transparent.

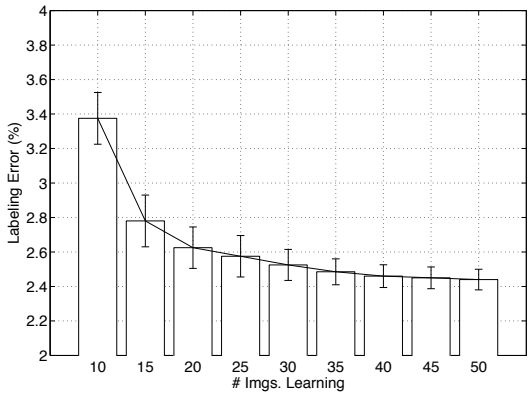

Figure 5. Variations in labelling errors with respect to the number of images used in the learning phase of the algorithm.

4.2. Periocular Biometrics Performance

To exemplify the usefulness of periocular segmentation algorithms, one *all-against-all* matching experiment was designed, using the method of Park *et* al. [13] and two different strategies to define the ROI: as baseline, the iris center was the unique reference for the ROI (upper-left image in Fig. 6). Next, according to the labels provided by the MRF, the center of mass of the cornea was used to define the ROI, which is obviously less sensitive to changes in gaze. Also, regions labelled as hair and glasses were disregarded from the recognition phase, considering that they likely suffer of significant variations among samples of a subject (upper-right image in Fig. 6). The Receiver Operating Characteristic curves for both variants are compared in the bottom plot of Fig. 6 and turn evident the benefits attained due to data segmentation (Equal error rate of 0.128 for the classical ROIs and 0.095 for the improved ROIs configuration). The improvements were substantial in all regions of the performance space, having at some operating points increased the system sensitivity over 10%. It should be stressed that no particular concerns were taken in optimizing the recognition method for the used data set, meaning that the focus was putted much more in the performance gap between both recognition schemes than in the recognition errors in absolute values, which are out of the scope of this paper.

5. Conclusions and Further Work

In this paper we have proposed an algorithm for *one-shot* labelling of all the components in the periocular region: iris, sclera, eyelashes, eyebrows, hair, skin and glasses. Our solution is composed of two major phases: 1) a group of local classification models gives the posterior probabilities for each pixel and class considered; 2) this *appearance*-based information is fused to geometrical constraints and shape priors to feed a two-layered MRF. One layer represents *pixels*, and analyzes the local data appearance while enforcing smoothness of the solutions. The second layer represents *components*, and assures that solutions are biologically plausible. By minimizing the MRF energy, the label of each pixel is found, yielding solutions that are robust against changes in scale, subjects' pose and gaze and dynamic lighting conditions.

As further directions for this work, our efforts are focused in estimate gaze / pose from the labelled data, in order to compensate for deviations before the recognition process.

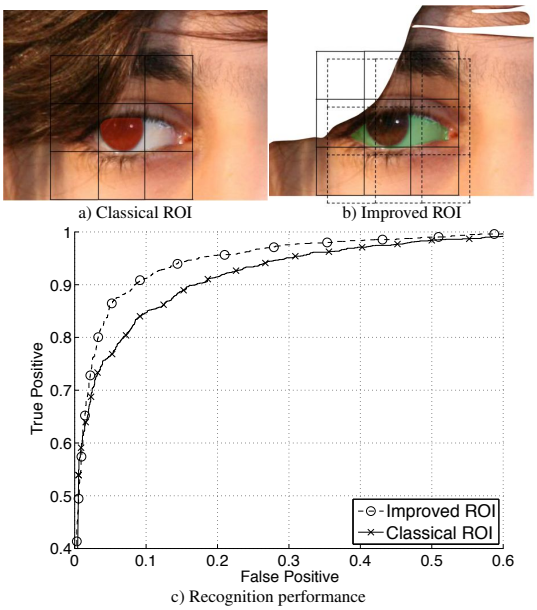

Figure 6. Improvements in periocular recognition performance due to the semantic categorization (labeling) of each pixel in the periocular region.

References

[1] A. Albiol, L. Torres and E. Delp. Optimum color spaces for skin detection. In Proceedings of the *International Conference on Image Processing*, vol. 1, pag. 122–124, 2001.

[2] C. Basca, M. Talos and R. Brad. Randomized Hough Transform for Ellipse Detection with Result Clustering. in Proceedings of the *The International Conference on Computer as a Tool EUROCON*, vol. 2, pag. 1397–1400, 2005.

[3] A. Besbes, N. Komodakis, G. Langs and N. Paragios. Shape Priors and Discrete MRFs for Knowledge-based Segmentation. in Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, pag. 1295–1302, 2009.

[4] S. Bharadwaj, H. Bhatt, M. Vatsa and R. Singh. Periocular biometrics: When iris recognition fails. In Proceedings of the 4^{th} *IEEE International Conference on Biometrics: Theory Applications and Systems*, pag. 1–6, 2010.

[5] B. Caputo, S. Bouattour and H. Niemann. Robust appearance-based object recognition using a fully connected Markov random field In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, vol., 3, pag. 565–568, 2002.

[6] S. Crihalmeanu and A. Ross. Multispectral scleral patterns for ocular biometric recognition. *Pattern Recognition Letters*, vol. 33, no. 14, pag. 1860–1869, 2012.

[7] P. Felzenszwalb and D. Huttenlocher. Efficient Belief Propagation for Early Vision. *International Journal of Computer Vision*, vol. 70, no. 1, pag. 41–54, 2006.

[8] B. Glocker, D. Zikic, N. Komodakis, N. Paragios and N. Navab. Linear Image Registration Through MRF Optimization. In Proceedings of the *IEEE International Symposium on Biomedical Imaging*, pag. 422–425, 2009.

[9] K. Oh, B-S. Oh, K-A. Toh, W-Y. Yau and H-L. Eng. Combining sclera and periocular features for multi-modal identity verification. *Neurocomputing*, vol. 128, pag. 185–198, 2014.

[10] Z. Kato and T.C. Pong. A Markov random field image segmentation model for textured images. *Image and Vision Computing*, vol. 24, pag. 1103–1114, 2006.

[11] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pag. 1568–1583, 2006.

[12] T. Ojala, M. PietikŁinen and D. Harwood. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In Proceedings of the 12^{th} *IAPR International Conference on Pattern Recognition*, pag. 582–585, 1994.

[13] U. Park, R. Jillela, A. Ross and A. Jain. Periocular Biometrics in the Visible Spectrum. *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pag. 96–106, 2011.

[14] U. Park, A. Ross and A. Jain. Periocular biometrics in the visible spectrum: A feasibility study. In Proceedings of the *IEEE* 3^{rd} *International Conference on Biometrics: Theory, Applications, and Systems*, pag. 1–6, 2009.

[15] H. Proença. Iris Recognition: On the Segmentation of Degraded Images Acquired in the Visible Wavelength. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pag. 1502–1516, 2010.

[16] U. Park, A. Ross and A. Jain. Matching highly non-ideal ocular images: an information fusion approach. In Proceedings of the *IEEE* 5^{th} *International Conference on Biometrics*, pag. 446–453, 2012.

[17] C-W. Tan and A. Kumar. Towards Online Iris and Periocular Recognition Under Relaxed Imaging Constraints. *IEEE Transactions on Image Processing*, vol. 22, no. 10, pag. 3751–3765, 2013.

[18] T. Tan, X. Zhang, Z. Sun and H. Zhang. Noisy iris image matching by using multiple cues. *Pattern Recognition Letters*, vol. 33, pag. 970–977, 2012.

[19] C. Wang, M. Gorce and N. Paragios. Segmentation, Ordering and Multi-Object Tracking using Graphical Models. In Proceedings of the *in* 12^{th} *International Conference on Computer Vision*, pag. 747–754, 2009.

[20] D. Woodard, S. Pundlik, P. Miller, R. Jillela and A. Ross. On the fusion of periocular and iris biometrics in non-ideal imagery. In Proceedings of the *in* 20^{th} *International Conference on Pattern Recognition*, pag. 201–204, 2010.

[21] D. Woodard, S. Pundlik, P. Miller and J. Lyle. Appearance-based periocular features in the context of face and non-ideal iris recognition. *Signal, Image and Video Processing*, vol. 5, pag. 443–455, 2011.