An Alternative Approach to Face and Iris Recognition Under Unconstrained Settings

João C. Monteiro and Jaime S. Cardoso
INESC TEC (formerly INESC Porto) and Faculdade de Engenharia, Universidade do Porto, Porto, Portugal
Email: joao.carlos.monteiro@fe.up.pt, jaime.cardoso@inescporto.pt

Abstract—In recent years, the focus of research in face and iris biometrics has been on the exploration of new traits to aid in the recognition process. The present work analyzes the periocular region as an alternative in less controlled acquisition scenarios. We based our work on the Universal Background Modelling framework proposed for speaker verification. SIFT keypoint descriptors were used as features and GMMs for modelling. The algorithm presented a rank-1 recognition rate of 81.79% and a decidability index of 2.99.

I. INTRODUCTION

Personal identification plays an important role in almost everyone’s daily activities. In the field of identification, biometrics works by recognizing specific patterns within highly unique biological traits. By attempting identification based on physiological and behavioural traits we are testing someone by who (s)he is, instead of relying on something (s)he owns or knows. Such approach seems likely to be the way forward.

Over the past few years face and iris have been on the spotlight of many research works in biometrics. The face is an easily acquirable trait with a high degree of uniqueness, while the iris, the colored part of the eye, presents unique textural patterns from its random morphogenesis [1]. These advantages, however, fall short when low-quality images are presented to the system. To overcome such limitations, the research focus turned to the study of alternative traits capable of supporting face and iris recognition under less ideal scenarios.

The periocular region, in the immediate vicinity of the eye as depicted on Figure 1, is one of such unique traits. In this work we present a new approach to periocular recognition, as an alternative to both iris and face recognition, under less ideal acquisition conditions.

![Fig. 1. Example of periocular regions from both eyes, extracted from a face image. Adapted from [2]](image)

II. PROPOSED METHODOLOGY

On the present work we aimed to achieve high performance periocular recognition by taking advantage of techniques with proven efficiency in other biometric applications. With that in mind we followed a Universal Background Modelling (UBM) framework, often found in the speaker recognition literature [3]. We aimed to assess the versatility of such approach, regardless of the tested biometric trait and the nature of the features chosen for its representation.

Universal background modelling is a popular strategy in the field of biometrics. Such popularity can be easily understood by stating the problem of biometric verification as a basic hypothesis test. Given a biometric sample Y and a claimed ID, S, we define:

\[ H_0: Y \text{ belongs to } S \]
\[ H_1: Y \text{ does not belong to } S \]

as the null and alternative hypothesis, respectively. The optimal decision is taken by a likelihood-ratio test:

\[
\frac{p(Y|H_0)}{p(Y|H_1)} \begin{cases} \geq \theta & \text{accept } H_0 \\ \leq \theta & \text{accept } H_1 \end{cases}
\]

where \( \theta \) is the decision threshold for accepting or rejecting \( H_0 \), and \( p(Y|H_i) \) is the likelihood of observing \( Y \) knowing that \( H_i \) is true. The goal of a biometric verification system can, thus, be accomplished by the computation of the likelihood values \( p(Y|H_0) \) and \( p(Y|H_1) \) for a given sample. It is intuitive to note that \( H_0 \) will correspond to a model \( \lambda_{hyp} \), that characterizes the hypothesized individual, while \( H_1 \) will represent the alternative hypothesis, that is, the model of all the alternatives to the hypothesized individual, \( \lambda_{hyp} \). The most common designation in literature for such a model is universal background model or UBM.

Defining an objective way of quantifying \( p(Y|H_0) \) and \( p(Y|H_1) \) is the true challenge when following this approach to biometric recognition. On the present work we explored Gaussian Mixture Models, which present both the robustness of parametric unimodal Gaussian density estimates, as well as the ability of non-parametric models to fit non-Gaussian data. We train our GMM models on a set of SIFT keypoints extracted from each periocular image.

For the UBM training we select a large subset of images, so as to faithfully cover a representative user space. SIFT keypoints are extracted from each image on the subset and used to model \( H_1 \). Individual specific models, \( H_1^{(i)} \), are then adapted from \( H_1 \) in a maximum a posteriori sense, using individual specific subsets of images and their respective SIFT
keypoints. The GMM adaptation process was first proposed by Reynolds for speaker verification, and details can be found in [3].

III. RESULTS

A. Experimental setup

The proposed algorithm was tested on a very noisy color iris database, the UBIRIS.v2 database. Images in UBIRIS.v2 were captured under non-constrained conditions (at-a-distance, on-the-move and on the visible wavelength), with corresponding realistic noise factors. Some example can be observed on Figure 2. A subset of the original database, composed by 3000 images from 100 distinct individuals was created. Per individual, 10 images were used to adapt the UBM and 20 to test performance. A random set of 1000 images was also built to train the UBM.

![Example images from the UBIRIS.v2 database.](image)

Fig. 2. Example images from the UBIRIS.v2 database.

Performance was evaluated in both verification and identification functioning modes. Regarding the former we analyzed the equal error rate (EER) and the decidability index (DI). The EER is observed at the decision threshold, \( \theta \), where the errors corresponding to falsely accepting and falsely rejecting \( H_0 \) occur with equal frequency. The global behaviour of both types of errors is often analyzed through receiver operating characteristic (ROC) curves.

The DI quantifies the separation of the “genuine” and “impostor” likelihood score distributions, as follows:

\[
DI = \frac{|\mu_g - \mu_i|}{\sqrt{0.5(\sigma_g^2 + \sigma_i^2)}}
\]

where \((\mu_g, \sigma_g)\) and \((\mu_i, \sigma_i)\) are the mean and standard deviation of the genuine and impostor score distributions, respectively.

For identification we chose to analyze results using cumulative match curves (CMC). These curves represent the rate of correctly identified individuals, by checking if the true identity is present in the \( N \) highest ranked identities. The \( N \) parameter is generally referred to as rank. That allows us to define the rank-1 recognition rate as the value of the CMC at \( N = 1 \).

B. Recognition Performance

Regarding the performance of the proposed algorithm we analyzed the effect of the GMM order, \( M \), and observed a performance plateau at \( M = 64 \). As the UBIRIS.v2 database is composed of RGB images, we also explored the effect of color channel fusion. Table I summarizes the most relevant results. It is easily observable that there is no significant difference between the individual performance of the Red, Green and Blue channels. However by testing simple combinations of the individual channel results, we observed some interesting results. Fusion was performed at the score-level, i.e. the likelihood-ratio values \( \frac{p(Y | H_0)}{p(Y | H_1)} \) were computed for each singular channel and fusion was performed over the 3 resulting values. A few simple operations, namely the mean, maximum, minimum and median values of the 3 likelihood-ratios for individual channels, were tested.

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>Green</th>
<th>Blue</th>
<th>Mean</th>
<th>Median</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>EER</td>
<td>9.12</td>
<td>8.53</td>
<td>8.60</td>
<td>5.16</td>
<td>8.60</td>
<td>7.13</td>
<td>7.69</td>
</tr>
<tr>
<td>Decidability</td>
<td>2.46</td>
<td>2.61</td>
<td>2.57</td>
<td>2.99</td>
<td>2.57</td>
<td>2.69</td>
<td>2.63</td>
</tr>
<tr>
<td>Rank-1</td>
<td>64.83</td>
<td>69.23</td>
<td>65.98</td>
<td>81.79</td>
<td>73.14</td>
<td>70.89</td>
<td>72.44</td>
</tr>
</tbody>
</table>

TABLE I. EFFECT OF COLOR CHANNEL AND FUSION STRATEGY ON RECOGNITION PERFORMANCE.

The mean-based fusion presented significantly higher performance than any singular channel (39% better in EER, 15% better in decidability and 18% better in Rank-1 recognition rate) and any other fusion strategy (33% better in EER, 10% better in decidability and 11% better in Rank-1 recognition rate), even though it is a very simple approach. These results seem to point to a considerable advantage in exploring fusion of color channel information in order to improve recognition performance.

IV. Conclusion

The idea of adapting individual specific models from an universal model has been thoroughly documented and presented good results in the field of biometrics. The proposed algorithm extended the idea to periocular recognition, a newly explored alternative to face and iris biometrics in less constrained acquisition scenarios. The obtained results are still far from those acquired for both iris and face under controlled scenarios, but reveal some potential for future improvements.

It is very interesting to observe how versatile the idea of a UBM is. However we might wonder if we are not compromising better performance by not exploring more trait-specific approaches. Nevertheless, working with this framework opens the possibility of easily performing feature-level fusion and expand our studies to multimodal biometrics.

ACKNOWLEDGMENT

The first author would like to thank Fundação para a Ciência e Tecnologia (FCT) - Portugal the financial support for the PhD grant with reference SFRH/BD/87392/2012.

REFERENCES