Iris liveness detection methods in the mobile biometrics scenario

Ana F. Sequeira, Juliano Murari and Jaime S. Cardoso

Abstract—Biometric systems based on iris are vulnerable to direct attacks consisting on the presentation of a fake iris to the sensor (a printed or a contact lenses iris image, among others). The mobile biometrics scenario stresses the importance of assessing the security issues. The application of countermeasures against this type of attacking scheme is the problem addressed in the present paper. Widening a previous work, several state-of-the-art iris liveness detection methods were implemented and adapted to a less-constrained scenario. The proposed method combines a feature selection step prior to the use of state-of-the-art classifiers to perform the classification based upon the “best features”. Five well known existing databases for iris liveness purposes (Biosec, Clarkson, NotreDame and Warsaw) and a recently published database, MobBIOfake, with real and fake images captured in the mobile scenario were tested. The results obtained suggest that the automated segmentation step does not degrade significantly the results.

I. INTRODUCTION

Biometric systems rather identify an individual by what he is instead of based on something he knows or possesses. Considering that any piece of material or knowledge can be fraudulently acquired biometrics can offer several advantages over classical security methods. However, in spite of its advantages, biometric systems have some drawbacks, including the lack of secrecy, the fact that a biometric trait cannot be replaced and its vulnerability to external attacks which could decrease their level of security. It is necessary to keep in mind the security issues when we explore a whole new world of possibilities in this networked society of nowadays. Our mobile devices are turning into storages of personal, professional, commercial, and other kinds of information. This information is intended by the users to be kept confidential. The necessity of controlling the access to this information opens the way to the concept of “mobile biometrics”.

Among the different existing biometric traits, iris has been traditionally regarded as one of the most reliable and accurate. Also the imaging properties of the handheld devices makes this trait instinctive to use. Therefore the development of iris liveness detection techniques is crucial for the deployment of iris biometric applications in daily life.

In a previous work we explored the use of countermeasures against spoofing attacks [1]. These attacks may consist on presenting a synthetically generated iris to the sensor so that it is recognized as the legitimate user and access is granted.

The most common and simple approaches are those carried out with high quality iris printed images [2]. However, other more sophisticated threats have also been reported in the literature such as the use of contact lenses [3].

In the present work we lean upon the previous work [1] and broaden the application of the method pursuing more unconstrained conditions. The methods for liveness detection may be hardware or software based. In the present work we use a software based approach. This countermeasures are to prevent attacks at sensor level. So, the fake irises are detected once the sample has been acquired with a standard sensor. The key point of the process is to find a set of discriminant features which permits to build an appropriate classifier which gives the probability of the sample vitality given the extracted set of features. In the previous work, the segmentation of iris was done manually, however, this is unrealistic if we are aiming a real world application. To overcome this less realistic scenario, in the present work we applied a automatic segmentation method. Therefore some of the implemented state-of-the-art methods were adapted to the non-circular contours obtained by the segmentation method. For the feature extraction, several state-of-the-art methods were implemented to perform the feature extraction that will help to discriminate between fake and real iris. Like previously, the extracted features were then included in a feature selection framework so as to determine the best cardinality and the best subset that conducts to the highest classification rate. Three different classifiers were used: Discriminant Analysis, K Nearest Neighbours and Support Vector Machines. The implemented methods were tested in four well known existing databases for iris liveness purposes (Biosec, Clarkson, NotreDame and Warsaw) and in a recently published database, MobBIOfake, with real and fake images captured in the mobile imaging scenario.

II. IRIS LIVENESS DETECTION

The problem of liveness detection of a biometric trait can be seen as a two class classification problem where an input trait sample has to be assigned to one of two classes: real or fake. The key point of the process is to find a set of discriminant features which permits to build an appropriate classifier which gives the probability of the sample vitality given the extracted set of features [4].

Biometric recognition systems are vulnerable to be spoofed by fake copies [5], for instance, fake finger tips made of commonly available materials such as clay and gelatine. Iris is no exception. There are potential threats for iris-based systems, the main are [6]:

- Eye image: Screen image, Photograph, Paper print, Video signal.
The feasibility of some attacks have been reported by some researchers [7], [5], [8] who showed that it is actually possible to spoof some iris recognition systems with printed iris and well-made color iris lens. Therefore, it is important to detect the fake iris as much as possible [6].

Several liveness detection methods have been presented through the past recent years. In fact, anti-spoofing techniques were presented that use physiological properties to distinguish between real and fake biometric traits. This is done in order to improve the robustness of the system against direct attacks and to increase the security level offered to the final user. Iris liveness detection approaches can broadly be divided into: i) software-based techniques, in which the fake irises are detected once the sample has been acquired with a standard sensor (i.e., features used to distinguish between real and fake eyes are extracted from the iris image, and not from the eye itself), and ii) hardware-based techniques, in which some specific device is added to the sensor in order to detect particular properties of a living iris such as the eye heppus (which is the permanent oscillation that the eye pupil presents even under uniform lighting conditions) or the pupil response to a sudden lighting event (e.g., switching on a diode) [4]. According to this author, even though hardware-based approaches usually present a higher detection rate, the software-based techniques have the advantage of being less expensive (as no extra device in needed), and less intrusive for the user (very important characteristic for a practical liveness detection solution). In general, a combination of both type of anti-spoofing schemes would be the most desirable approach to increase the security level of biometric systems. [4]

In this work we focus on software based techniques since these are more easily and affordable applicable in real-world applications.

In the literature we found that the methods of liveness detection may be classified into four categories based on the physical features of biometric and liveness data and the timing of measurement [9]. In this framework, the biometric data are used in the iris recognition and the liveness data are used in the liveness detection. We can itemize the four categories:

- **Perfect matching model**: Both biometric and liveness data are simultaneously obtained from the same physical feature.
- **Simultaneous measuring model**: Biometric and liveness data are simultaneously obtained from different physical features.
- **Same biometric measuring model**: Biometric and liveness data are obtained from the same physical feature with different timings.
- **Independent measuring model**: Biometric and liveness data are obtained from different features with different timings.

The ideal configuration of liveness detection for biometrics recognition is represented by the perfect matching model with the highest ability to distinguish between live and fake irises [10].

The potential of quality assessment to identify real and fake iris samples acquired from a high quality printed image has previously been explored as a way to detect spoofing attacks [4]. Some quality based features have been used individually for liveness detection in traits such as iris [10], [3] or face [11]. A strategy based on the combination of several quality related features has also been used for spoofing detection in fingerprint based recognition systems [12] as well as in iris liveness detection [4]. In this latter work, a set of quality measures are used as iris liveness detection features to aid the classification of fake or real iris images included in a framework of feature selection. We find in literature that works concerning the quality of iris images are often the starting point to iris liveness detection techniques. One example is the assessment of the iris image quality based on measures like occlusion, contrast, focus and angular deformation [13], other is the use of texture analysis of the iris [14], among others like, for example, the analysis of frequency distribution rates of some specific regions of iris [15].

The way forward seems to be the development of techniques for iris liveness detection that work well independently of the particular characteristics of the databases available nowadays. It is required to develop and improve methods as well as to construct new databases in less constrained conditions.

### III. METHODS

In the previous related work [1] we choose to make the segmentation process manually, in order to ensure reasonable accuracy. In the present work we move a step forward in pursuing an automated method suitable to real world applications. For the feature extraction we implemented five different methods to extract thirteen features. Two of the methods use the entire eye image while the others use only the iris region.

#### A. Robust segmentation under unconstrained scenarios

The segmentation algorithm, performs automatic detection of the inner and outer contours of the iris. The method focuses on mutual context information from iris center and iris limbic and pupillary contours to perform robust and accurate iris segmentation in noisy images.

The simultaneous detection of the iris center and limbic contour are addressed by first over-detecting center candidates, followed by a contour detection around each of them. The center candidates are estimated using a convergence index filter methodology [16]. Next, a window centered in each candidate is converted into the polar domain followed by shortest path algorithm to determine good closed paths around the center. Using combined data from the center and
respective contour, the best pair center/contour is selected. Finally, the pupillary segmentation is performed using a new polar image around the centroid of the detected limbic contour. For more details on the method see [17].

In the following sections we present the methods used for feature extraction.

B. Algorithm 1 - High Frequency Power

The High Frequency Power algorithm, which provides feature 1, works on the whole image and measures the energy concentration in the high frequency components of the spectrum using a high pass convolution kernel of 8x8. The application of this convolution is a good Fourier Transform approximation and works as high frequency spectral analysis, which can be considered an estimator of focus [18]. The focus of a real iris, as it is a 3D volume, is different from a fake iris focus, which has a 2D surface. For more details on the method see [4].

C. Algorithm 2 - Local Contrast

The Local Contrast algorithm, which provides feature 2, is based on the bounding box that involves the iris and the pupil. The bounding box is divided in blocks of $P \times P$ and for each block it is applied the Fast Fourier Transform (FFT) algorithm to extract the medium power frequencies, which better represents the contrast. The final value is given by the number of blocks with medium values (between 20 and 60) divided by the total number of blocks. This algorithm was inspired in an occlusion estimation technique [13] and it was adapted for contrast estimation for iris liveness detection in [4] where more details can be found.

D. Algorithm 3 - Global Contrast

The Global Contrast algorithm, which provides feature 3, explores the fact that parts extremely bright or dark of the image are not useful and can be considered as noise. Thus, pixels near medium value (128 in 8-bit image) are considered of best contrast [13]. In order to quantify the contrast, the original pixels values are normalized between 0 and 25 according to a normalization function. Original pixels near medium value will get higher values in the normalized scale, as well as very low and very high values ($<10$ and $>245$) are normalized to 0. This measure was presented in [13] and it was adapted for global contrast estimation for iris liveness detection in [4] where more details can be found.

E. Algorithm 4 - Adapted Frequency Distribution Rates

The Adapted Frequency Distribution Rates algorithm consists in different mathematical combinations of three different parameters which consider respectively the power of the low ($F_1$), medium ($F_2$), and high ($F_3$) frequencies (computed according to the 2D Fourier Spectrum) from two iris subregions. In the original method the two regions were determined by the circular contours of iris and placed in the horizontal direction. In this adaptation the position and radius of the two regions may differ due to the non-circular iris contours provided by the segmentation process. This is illustrated in Figure 1. Then each subregion is divided in three circular concentric regions determining the three values: $F_1^i$ (central circle), $F_2^i$ (middle circular ring) and $F_3^i$ (outer circular ring), for $i = 1, 2$ (corresponding to the left and right regions), as depicted in Figure III-E. The final values $F_1$, $F_2$ and $F_3$ are given by averaging the two regions. More details on the original method can be found in [15], [4]. With the three final frequencies we extract six different combinations, represented in Table I.

\[ \begin{array}{c|c}
\text{Features no.} & \text{Combination} \\
\hline
4 & F_1 + F_2 + F_3 \\
5 & F_2/(F_1 + F_3) \\
6 & F_2 \\
7 & F_3 \\
8 & (F_1 + F_2)/F_3 \\
9 & (F_1 \times F_2)/F_3 \\
\end{array} \]

F. Algorithm 5 - Statistical Texture Analysis

The Statistical Texture Analysis algorithm was developed as a contact lens countermeasure. The outer portion of the color contact lens (corresponding to regions closer to outer circle) provides the most useful texture information for fake iris detection since this section of the fake iris is insensitive to the pupil dilution [14]. The region of interest is the lower part of the iris in order to minimize the occlusion by the eyelashes and eyelids, which in general occurs in the upper iris portion. In order to achieve invariance to translation and scale, the region of interest is further normalized to a rectangular block of a fixed size $W \times H$.

Then GLCM (Gray Level Co-occurrence Matrix) [19], is calculated and four features are extracted: the mean ($\mu$) and
standard deviation (σ), direct from the normalized region of interest, and the contrast (con) and the energy (e) from the GLCM matrix. These measures will provide features 10 to 13 and for more details on the calculation of its values see [14], [1].

G. Feature Selection

The algorithms implemented originated 13 different features. It is convenient to search for the optimum number and set of features. To exhaustively test all possibilities is not feasible. Therefore we use the “Sequential Forward Floating Selection” (SFFS) [20] to perform feature selection. The SFFS is basically a combination of search methods such as “Plus-l-Minus-r” [21] and Sequential Forward Search (SFS) [22]. The appearance of “floating” comes from the fact that the values l and r are not fixed, i.e., they can “float”. Another aspect is the dominant direction of search, including (forward) or excluding (backward) characteristics [20]. We use the Mahalanobis distance as criterion function. The SFFS has shown to be competitive when compared to other selection techniques [23].

IV. EXPERIMENTAL SETUP

A. Datasets

The implemented methods were tested in five databases for iris liveness detection evaluation: four well know existing datasets (referred as Biosec, Clarkson, NotreDame and Warsaw, for simplicity) and a recently published one (MobBIOfake).

1) Biosec (C): The Biosec [24] database comprises a total of 1600 images: 800 real images and its corresponding 800 fake samples. All images are in greyscale and its dimensions are 640 x 480 [4]. The two eyes of the same individual are considered as different users.

2) Clarkson (C): The subset of Clarkson database that we used was made available under request to participants of the LivDet-2013 competition [25]) and contains 270 real iris images and 400 fake iris images. The fake samples are images of eyes with contact lenses comprising 14 models of contact lenses.

3) NotreDame (ND): The subset of images from the ND_Cosmetic_Contact_Lenses_2013 Dataset [26] that we used was made available under request to participants of the LivDet-2013 competition [25]). This dataset contains iris images of subjects without contact lenses, with soft contact lenses, and with cosmetic contact lenses, acquired using an LG 4000 iris sensor. The subset comprises 2000 “real” iris images and 1000 “fake” images.

4) Warsaw (W): The subset of Warsaw database [27] that we used was made available under request to participants of the LivDet-2013 competition [25]) and contains 228 real images and 203 fake images. The fake samples are printed images of real iris images.

5) MobBIOfake (Mf): The MobBIOfake [1] database is comprised of 800 iris images and its corresponding fake copies, captured with the same portable device and in similar conditions. This database is comprised of 800 real iris images and its corresponding 800 fake printed images. This database was constructed upon the MobBIO Multimodal Database [28] which comprised samples of voice, face and iris captured with a handheld device.

B. Methodology

The proposed method is depicted in Figure 2.

![Fig. 2. Steps of the proposed method](image)

The first step of the method is the segmentation. We used a automatic process described in section III-A.

The second step was the feature extraction. This comprises the application of the methods described in subsections III-B, III-C, III-D, III-E and III-F.

The next step was the feature selection. This comprises the application of the method Sequential Forward Floating Search, described in subsection III-G, before applying the classifiers to evaluate the proposed methods. We ran the SFFS to obtain the best subset for each cardinality from $N = 2$ to $N = 12$ features.

The last step was the classification. We used three state-of-the-art classifiers: Discriminant Analysis (DA), k-Nearest Neighbours (kNN) and Support Vector Machines (SVM). For each cardinality, ($N = 2, ... , 12$), the results of classification were obtained calculating the average of the results of 50 runs for classification of the images based on the corresponding best $\hat{N}$ features. The results were obtained by randomly dividing, in each run, the total samples in the training (62.5%) and testing (37.5%) sets. The parameter $k$ in kNN was optimized using cross-validation, and tested in the interval [1, 20] by steps of 1. For the SVM, we used a polynomial kernel and also used cross-validation for optimization of the parameters. It was performing a ”grid-search” on the parameters of the models. Exponentially growing sequences of $C$ were tested: $C = 2^N$ with $N$ varying between −1 and 15. For the polynomial degree, $d$, values tested were: $d = 1, 2, 3, 4, 5$.

For the evaluation of the accuracy of the features extracted in discriminating between fake and real images, we used the Equal Error Rate (EER). The EER is obtained when the false acceptance rate (FAR) and the False Rejection Rate (FRR) are equal. For the classification results we use the missclassification rate averaged over the 50 runs.
C. Experimental Results and Discussion

In this section we present the results obtained by the proposed method for iris liveness detection.

1) Automated Segmentation: Even though we aim to automatize our method the fact is that the segmentation step is by itself quite challenging. Although it is possible to find in the literature methods with good performance rates there is still much to improve. Specially concerning more unconstrained imaging scenarios. Due to the variable nature of images tested, the segmentation results leads to better or worse results depending on the database. It is necessary to balance the two aspects: automation versus quality. By one side, we intended to use an automatic segmentation process, but, on the other side, we wanted to avoid the excess of errors due to bad performed segmentation. Therefore, we choose to reject some of the segmented images in order to guaranty a minimum of accuracy of the method. It is almost as difficult to automatically decide about the quality of a result of a segmentation method as to perform the segmentation. This selection was made manually. Some examples are depicted in Figure 3.

![Examples of contours obtained in the segmentation step.](image)

In Table II it is shown the number of images accepted for each database.

<table>
<thead>
<tr>
<th>DB</th>
<th>Number of images</th>
<th>DA</th>
<th>C</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>real</td>
<td>fake</td>
<td>μ</td>
<td>σ</td>
</tr>
<tr>
<td>B</td>
<td>686</td>
<td>524</td>
<td>9.0</td>
<td>1.1</td>
</tr>
<tr>
<td>C</td>
<td>205</td>
<td>311</td>
<td>5.8</td>
<td>1.4</td>
</tr>
<tr>
<td>Mf</td>
<td>717</td>
<td>731</td>
<td>20.9</td>
<td>1.4</td>
</tr>
<tr>
<td>ND</td>
<td>1709</td>
<td>805</td>
<td>34.6</td>
<td>1.5</td>
</tr>
<tr>
<td>W</td>
<td>209</td>
<td>182</td>
<td>21.6</td>
<td>2.8</td>
</tr>
</tbody>
</table>

2) Methods not requiring contour detection: In Table III we show the results of classification for each database using only three features. The three features used do not require the detection of both, pupillary and limbic, contours. The algorithms 1 and 3 deal with the complete image. Algorithm 2 use a bounding box around the iris region. We included this method in this experiment taking in account that is much simpler to obtain a bounding box than performing the detection of both iris contours.

<table>
<thead>
<tr>
<th>DB</th>
<th>DA</th>
<th>C</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μ</td>
<td>σ</td>
<td>μ</td>
</tr>
<tr>
<td>B</td>
<td>0.8</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>C</td>
<td>9.1</td>
<td>2.1</td>
<td>6.2</td>
</tr>
<tr>
<td>Mf</td>
<td>16.9</td>
<td>1.6</td>
<td>16.3</td>
</tr>
<tr>
<td>ND</td>
<td>12.0</td>
<td>0.9</td>
<td>12.4</td>
</tr>
<tr>
<td>W</td>
<td>10.5</td>
<td>2.4</td>
<td>13.6</td>
</tr>
</tbody>
</table>

Concerning the methods that does not require iris contour detection, the best result was obtained for the Clarkson database. Observing the results we note that in this case the error rates are rather high.

3) Methods requiring contour detection: In this section we present the results of the proposed method (comprising algorithms 1 to 5) in the two different segmentation scenarios: manual and automated segmentation.

In Table IV we show the results of classification for each database using the information from manual segmentation. The sets of images we used to obtain these results were the subset of images selected after the segmentation step. This was made so that we can compare the results of the two scenarios.

<table>
<thead>
<tr>
<th>DB</th>
<th>DA</th>
<th>kNN</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μ</td>
<td>σ</td>
<td>N</td>
</tr>
<tr>
<td>B</td>
<td>0.5</td>
<td>0.3</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>6.2</td>
<td>1.6</td>
<td>11</td>
</tr>
<tr>
<td>Mf</td>
<td>7.7</td>
<td>0.7</td>
<td>9</td>
</tr>
</tbody>
</table>

The best result (for the manual method) was obtained for the Biosec database, 0.5%. The worse result was obtained for MobBIOfake, 14.7%. This confirms the results observed in a previous work [1] in which we were able to conclude that the MobBIOfake database was more challenging than the existing databases for iris liveness detection tested. In the referred work the set of databases was smaller, comprising the Biose, Clarkson and MobBIOfake databases.

In Table V we show the results of classification for each database using the automatic segmentation.

The best result (for the automatic method) was obtained for the Biosec database, 2.0%, as for the manual method. Again, the worse result was obtained for MobBIOfake, 11.4%. This latter is not surprising since, as referred, we have observed...
in a previous work that this database was more challenging than the existing databases for iris liveness detection.

When we compare the two scenarios, we observe that for most databases the results did not significantly worsened moving from the manual to the automatic segmentation. In some situations, the error rate even lowered. This fact is not easily explained with the current experiments is requires further testing but we may consider some over fitting phenomena.

Considering the cardinalities of the best classification results, we observe that the best results are in most of the cases obtained with a high number of features. In our point of view this fact does not invalidate the usefulness of the feature selection although it may suggest to extend the set of features extracted.

Considering the classification methods, we observe that the SVM provided the best results when compared to the other classifiers: DA and kNN (with only one exception in ten results).

V. CONCLUSIONS AND FUTURE WORK

In this networked society of ours, mobile handheld devices have evolved from simple communication devices to mobile personal computers. Therefore, important personal private data as well as business data are stored on those mobile handheld systems. Naturally there is a strong need for user identification and access control. These need leads to the emerging field so called Mobile Biometrics. The current trend is not only to increase the recognition performance in mobile scenarios but also to increase the security of this processes. In this scenario the actuality of the iris liveness detection topic is unquestionable.

In this work, we compared four well known databases for iris liveness detection and a recently published database for iris liveness detection purposes with images acquired in unconstrained conditions and with a handheld device. The proposed method comprised state-of-the-art methods for iris liveness detection in a feature selection framework. Two different scenarios were tested whether the segmentation step was done manually or by an automated process. This latter proved to raise more difficulties to the liveness classification problem as the results worsened. In both scenarios the MobBIOfake database proved itself to be more challenging than the others.

Published works present methods tested with existing databases which achieve excellent results, (0% error classification rate). However, we note that some of these methods are closely connected with the particular database characteristics. In the present work we avoid measures such as ratios of pupil and iris radius or areas, for example. Also we introduce a new variation which affects the overall results: the automatic segmentation step. Comparing our two scenarios, is not certain that automatic segmentation is a source of error for the liveness detection problem, only in some databases the classification results worsened. Anyway, we consider that this is the path to walk if we aim to pursue real-world applications in which segmentation must be done automatically and in real time.

For future work, we foresee the necessity of improving the existing methods for iris liveness detection and develop new ones more suitable to the new imaging scenarios.

REFERENCES