Pattern Recognition Letters
journal homepage: www.elsevier.com

Multimodal authentication on Smartphones: combining iris and sensor recognition for a double check of user identity.

Chiara Galdi[a,] ∗, Michele Nappi[b] and Jean-Luc Dugelay[a]

[a]*EURECOM, 450 Route des Chappes, CS 50193 - 06904 Biot Sophia Antipolis cedex, France*
[b]*Università degli Studi di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italia*

ABSTRACT

Iris recognition on mobile devices is a challenging task, performing acquisition via the embedded sensors can introduce the sensor interoperability problem. Biometric systems developed so far are limited in their ability of comparing biometric data originated by different sensors because they operate under the assumption that the data to be compared are obtained using the same sensor. This problem leaded to the development of biometric recognition algorithms able to work independently from the data source. In this paper, we get around the sensor interoperability problem leveraging on the picture differences due to acquisition by different sensors. We present a novel system that combines the recognition of user‘s iris and user’s device, i.e. something the user is plus something the user has. To do so, we adopted an iris recognition algorithm, namely Cumulative Sums, and a well-known technique in the image forensic field for camera source identification based on the extraction of the Sensor Pattern Noise. The two identification processes are performed on the same picture leading to a system with a good trade-off between ease of use and accuracy. The approach is tested on MICHE, a database composed by iris images captured with different mobile devices in unconstrained acquisition conditions.

Keywords: sensor recognition; iris recognition; multimodal system; smartphone; sensor interoperability; MICHE.

∗ Corresponding author. Tel.: +33-(0)4-93-00-81-67; fax: +33-(0)4-93-00-82-00; e-mail: galdi@eurecom.fr

1. Introduction

Biometric recognition for a long time has been used in confined spaces, usually indoor, where security-critical operation required high accuracy recognition systems, e.g. in police stations, banks, companies, airports (usually for frequent flyers, so just for a limited number of voyagers). Field activities, on the contrary, required more portability and flexibility leading to the development of devices for less constrained biometric traits acquisition and consequently of robust algorithms for biometric recognition in less constrained conditions [18]. However, the application of "portable" biometric recognition, was still limited in specific fields e.g. for immigration control, and still required dedicated devices.

A further step would be to spread the use of biometric recognition on personal devices, as personal computers, tablets and smartphones. Some attempts in this direction were made embedding fingerprint scanners in laptops or smartphones [26]. However, so far biometric recognition on personal devices has been employed just for a limited set of tasks, as to unlock the screen using fingerprints instead of passwords, PINs, or patterns. One of the reasons is that systems presented so far can be easily spoofed, as the well-known hacking of the Touch ID on iPhone6.

To improve biometric recognition robustness against attacks it is worth using multimodal recognition. This topic has already been addressed (see section 2) and few works suggest to combine two biometric traits for user identification on mobile devices, i.e. multi-biometric recognition. In this paper indeed, a multimodal authentication system that combines a biometric trait, namely the iris, with a personal object owned by the user, namely the smartphone is presented. This approach has several advantages:

- The genuine sample consists in the couple user-device, making more difficult the spoofing process;
- The two recognition processes are applied on a single photo of the eye captured by the user with his/her smartphone;
- Good trade-off between accuracy and ease of use;
- Performances of iris recognition and, in particular, of sensor recognition, are very high.

The system we propose is therefore a multimodal recognition system based on the combination of sensor recognition (hardwaremetry) and iris recognition (biometry), i.e. something the use has + something the user is. If we analyze the authentication systems security levels shown in Figure 1, we can see that the degree of security assured by the combination of biometry and a physical object is higher with respect of the use of biometry only [1]. The second aspect that we address in this paper is the sensor interoperability problem [28]. This problem rises up when the data to be compared (e.g. the pictures of the eye) are acquired with different sensors and thus contain differences depending on the sensor characteristics. As we will show later in the paper, this can affect the biometric algorithm performances since two pictures of the same eye can appear different because if they were captured by different devices. Our approach can be seen as a way to bypass the sensor interoperability problem, instead of focusing on the development of an algorithm able to operate regardless of the sensor employed, we leverage on the differences introduced by different sensors on photos in order to obtain a more robust recognition system.

We studied different techniques for the fusion of iris and sensor recognition. We tested the system on a database that accurately simulates the application of the system in real life. In fact MICHE [3] is an iris images database consisting of photos of the eyes of 75 different persons, captured with different mobile devices in different illumination conditions. Thanks to this database, it is possible to actually perform the double check of iris and device identity on a single photo and at once. Finally, we analyzed system response to noise, showing that the use of two traits (iris texture and Sensor Pattern Noise) of a so different nature can mitigate the deterioration of the global system performance, i.e. the more the quality of acquired iris images degrades the more the SPN is important in a verification process.

The outline of the paper is as follows: in chapter 2 we will present previous works on biometric recognition on mobile phones; the techniques employed for sensor recognition, iris recognition, score normalization, and fusion are presented in chapter 3, 4, 5, and 6 respectively; in chapter 7 experimental settings and results are illustrated, performance are assessed in terms of Cumulative Match Score curve (CMS), and Receiver Operating Characteristic curve (ROC); the paper ends in chapter 8 that summarize our main findings and conclusions.

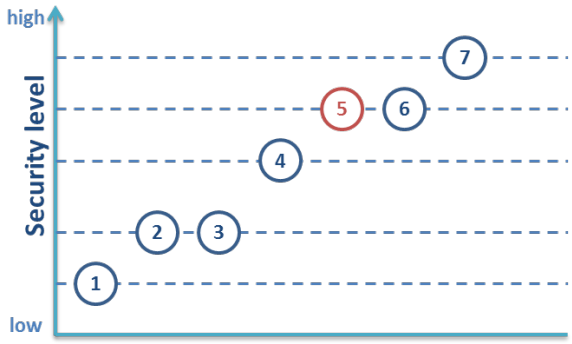

Figure 1 Authentication systems security levels: (1) Something the user knows; (2) Something the user has; (3) Something the user knows + something the user has; (4) Something the user is or does; (5) Something the user has + something the user is or does; (6) Something the user knows + something the user is or does; (7) Something the user knows + something the user has + something the user is or does.

2. Background

The biometric trait firstly chosen for biometric recognition on mobile phones, leveraging the presence of embedded cameras, is of course the face. In fact face recognition algorithms do not require high resolution images, and for this reason face was more suitable than iris at the beginning, when the resolution provided by mobile phone embedded cameras was limited. Some example of face recognition on mobile phone are presented in [4] and [5], the latter also addresses the problem of performing complex face recognition tasks on a mobile terminal. This could shorten the battery lifetime, while it is better to use the mobile phone only as an interface and perform all computationally heavy operations on the server side. In [6] the face recognition system presented also addresses the issue of using biometric recognition for security-critical operations, e.g. home banking, providing an anti-spoofing module and the opportunity of performing continuous recognition.

Nowadays smartphones provide built-in high resolution imaging sensors that can be used to perform also iris recognition. This gave the researchers the green light to study proper solutions to perform all the phases of iris recognition on mobile phones. For what concerns iris detection, in [7] and [8] methods for pupil and iris boundaries detection are presented, in these two works however, the databases employed were collected respectively

with a Samsung SPH-S2300 and Samsung SPH-2300 [9] (in [7] only 132 images were captured with the mobile phone and the others were from CASIA database [10]) which embed a 3.2 megapixel digital camera with a 3X optical zoom, which is a very specific imaging sensor that cannot commonly be found in the most popular smartphones. Toward the aim of providing a solution suitable for any kind of mobile devices, in [11] and [12] a database acquired with different mobile devices, namely MICHE database [3], is employed to test iris segmentation.

One of the first works investigating the possibility of optimizing iris segmentation and recognition on mobile phones is [13], Jeong *et al.* propose a method for computing the iris code based on Adaptive Gabor Filter. In [14], Park *et al.* present a recognition method based on corneal specular reflections, while Kang in [15] presents a method to pre-process iris in order to remove the noise related to occlusions of eyelids and improve system performances. In [16] and [17] an iris recognition system based on Spatial Histograms is presented. Finally, we focus on a work that represents a step forward the development of a secure authentication system via mobile phone, in fact in [18], De Marsico *et al.* present a face and iris recognition system that also integrates an anti-spoofing module.

3. Hardwaremetry

In order to recognize the sensor that captured a given photo, we implemented the Enhanced Sensor Pattern Noise (ESPN) based algorithm presented by Li in [19]. This method extracts from a picture the noise pattern of the sensor, it can also be used to distinguish cameras of the same model [20][21]. The approach presented by Li, is based on a previous work by Lukás et al. [20] in which the authors present the algorithm for extracting the Sensor Pattern Noise (SPN).

The ESPN is extracted from the Sensor Pattern Noise (SPN) by applying a filter that removes the details of the image located in the highest frequencies. The SPN is obtained using the following formula:

$$n=DWT(I)-F(DWT(I))$$

where *DWT()* is the discrete wavelet transform to be applied on image *I* and *F()* is a denoising function applied in the DWT domain. For *F()* we used the filter proposed in appendix A of [20]. In Figure 2 the denoising process is illustrated: Figure 2 (a) shows a sample of the MICHE database. We selected this image because it contains many "strong details", e.g. the frame of the glasses or the dark hair on a light background, that in the frequency domain are located in the high frequencies and can affect the Sensor Patter Noise extraction process. In Figure 2 (c) and Figure 2 (e), we can see how the denoising process has mitigated the presence of those details.

SPN is then enhanced as suggested in [19] according to the following formula:

$$n_e(i,j) = \begin{cases} e^{-0.5n^2(i,j)/\alpha^2}, & \text{if } 0 \le n(i,j) \\ -e^{-0.5n^2(i,j)/\alpha^2}, & \text{otherwise} \end{cases}$$

where n_e is the ESPN, n is the SPN, i and j are the indices of the components of n and n_e, and α is a parameter that we set to 7, as indicated in [19]. An example of the difference between the SPN and the ESPN is shown in Figure 3, the original picture contains many details that influence the SPN but they can be mitigated by the enhancing step.

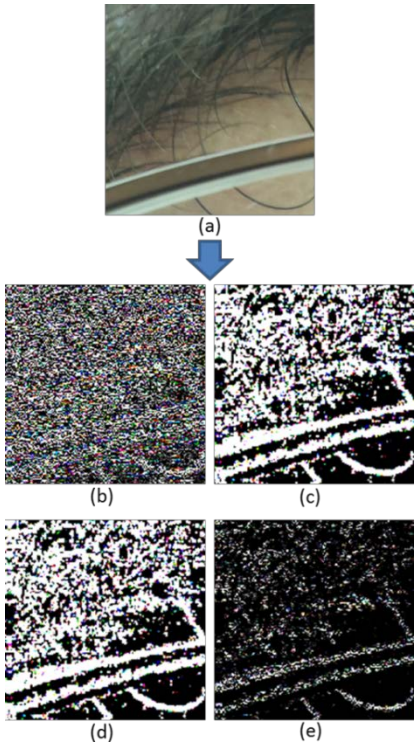

Figure 2 Denoising process: (a) original image; (b) original wavelet coefficients; (c) local variance; (d) selection of the minimum variance; (e) denoised wavelet coefficients.

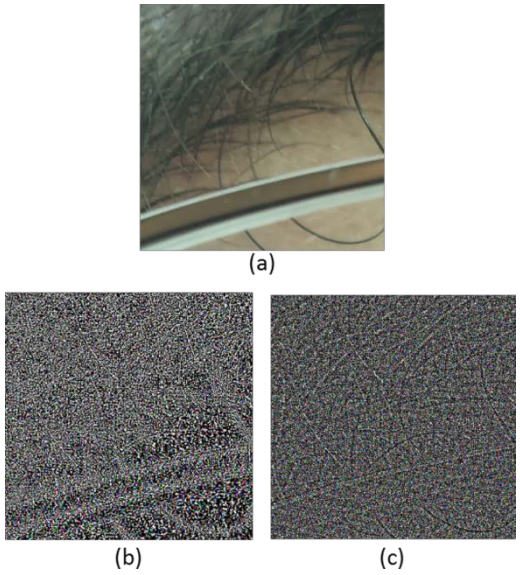

Figure 3 Sensor Pattern Noise enhancing: (a) original image; (b) the SPN extracted from the image contains image details (e.g. hairs, part of the eyeglasses frame); (c) ENSP, after the enhancing step the influence of image details is mitigated.

The process shown above, allows us to obtain the ESPN, i.e. the "fingerprint" of the sensor that captured the given photo. To associate then the extracted fingerprint to the correct sensor, we have to compare this fingerprint to the Reference Sensor Pattern Noise (RSPN) of the sensor. To extract the RSPN n_r of a sensor, we compute the average SPN over N photos acquired with the given camera (see section 7.2 for details):

$$n_r = \frac{1}{N} * \sum_{k=1}^{N} n_k$$

Where n_k is the SPN extracted from the k^{th} image. To compare the ESPN extracted from a photo with a RSPN of a sensor, we compute their correlation as follows:

$$\mathrm{corr}(\mathrm{n_e}, \mathrm{n_r}) = \frac{(\mathrm{n_e} - \overline{\mathrm{n_e}}) * (\mathrm{n_r} - \overline{\mathrm{n_r}})}{\|\mathrm{n_e} - \overline{\mathrm{n_e}}\| \|\mathrm{n_r} - \overline{\mathrm{n_r}}\|}$$

where the bar above a symbol denotes the mean value.

4. Biometry

Iris recognition on mobile devices is a challenging task, in fact, with respect to other dedicated iris acquisition devices, the smartphone embedded sensors introduce a number of noisy factor during the iris acquisition process [25]: out-of-focus, off-angle iris, rotated iris images, motion blurring, occlusions due to eyelashes, occlusions due to eyelids, occlusions due to eyeglasses, occlusions due to contact lenses, specular reflections, diffuse reflections, partially captured iris.

For this reason in this context it is preferable to adopt an iris recognition algorithm suitable for low quality iris images. The algorithm we employed is the Cumulative SUMs (CSUM) [22]. This method analyzes the image local variation in gray level. In our implementation, the iris image is first normalized transforming the Cartesian coordinates in polar ones, obtaining a rectangular shape. Then the image is subdivided into cells and for each cell, the representative value X is computed as the average gray level. Then the cells are grouped (horizontally and vertically in turn) and the average value $\overline{\mathrm{X}}$ of the representatives of the cells of each group is computed. An illustration of an iris image subdivided in cells and groups is shown in Figure 4. The cumulative sums are then computed over each group as follows:

$$\mathrm{S_0} = 0$$

$$\mathrm{S_i} = \mathrm{S_{i-1}} + (\mathrm{X_i} - \overline{\mathrm{X}}) \qquad \text{for } \mathrm{i} = 1, 2, \ldots, \mathrm{N}$$

where N is the number of elements of the group.

Finally, the iris code is generated comparing each pair of consecutive sums and assigning values 1 or 2 to a cell if the value of the corresponding sum contributes respectively to an upward slope or to a downward slope. Otherwise, value 0 is assigned to the cell.

The matching of the iris codes computed as explained before, is performed by Hamming distance.

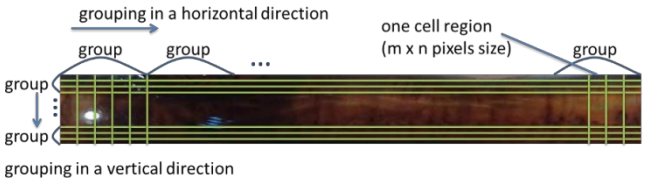

Figure 4 Cumulative Sums algorithm illustration

5. Score Normalization

Score normalization is a necessary step when combining different modules. The algorithms employed by each module can generate scores that are different in terms of distribution and numerical range. In the past several different methods of score normalization have been proposed [29], addressing different issues that can emerge during the fusion process. In our experiments, we tested five different normalization techniques, namely: Max-Min, Z-score, Median/MAD, TanH, and Sigmoidal. We will briefly explain these techniques in the following. Let's denote the set of K scores as: $S \rightarrow \{s_k\}, k = 1, 2, \ldots, K$, and the resulting set of normalized scores as: $S' \rightarrow \{s'_k\}, k = 1, 2, \ldots, K$.

5.1. Max-Min normalization technique

With the Max-Min technique, the scores are normalized based on the maximum and the minimum values in the scores set. The advantage of this simple technique is that the resulting scores set has a fixed numerical range: [0, 1]. In addition the shape of the original scores distribution is preserved. The Max-Min normalization technique can be implemented using the following formula:

$$s'_k = \frac{s_k - min_k\, s}{max_k\, s - min_k\, s}$$

5.2. Z-score normalization technique

The Z-score normalization technique is based on the calculation of the arithmetic mean and the standard deviation of the scores set. Thus, the resulting normalized scores set has a mean of zero and a standard deviation of one. However, this technique does not assures that the resulting scores set has a common numerical range, and it can also be sensitive to the presence of outliers. The Z-score normalization technique can be implemented using the following formula:

$$s'_k = \frac{s_k - mean(s)}{mean(s) - std(s)}$$

5.3. Median/MAD normalization technique

This technique is based on the median and median absolute deviation (MAD) that are insensitive to outliers. The Median/MAD normalization technique can be implemented using the following formula:

$$s'_k = \frac{s_k - median(s)}{MAD}$$

where $MAD = median(|s_k - median(s)|)$. This technique too has the disadvantage that does not preserve the input distribution and does not transform the scores into a common numerical range.

5.4. TanH normalization technique

The tanH technique was introduced by Hampel et al. [30]. It is robust and highly efficient and the normalization formula is:

$$s'_k = \frac{1}{2}\left\{tanh\left(0.01\left(\frac{s_k - \mu_{GH}}{\sigma_{GH}}\right)\right) + 1\right\}$$

where μ_{GH} and σ_{GH} are the mean and standard deviation estimates, respectively, of the genuine score distribution as given by Hampel estimators [30]. An advantage of this method is that it is not sensitive to outliers.

5.5. Sigmoidal normalization technique

Cappelli et al. [31] adopted this technique in order to combine the scores of a multimodal biometric system based on the combination of different fingerprint classifiers. The normalized score can be obtained by the following double sigmoid function:

$$s'_k = \begin{cases} \dfrac{1}{1+\exp\left(-2\left(\frac{s_k - t}{r_1}\right)\right)} & if\ s_k < t, \\ \dfrac{1}{1+\exp\left(-2\left(\frac{s_k - t}{r_2}\right)\right)} & otherwise, \end{cases}$$

where t is the reference operating point and r_1 and r_2 denote the left and right edges of the region in which the function is linear, i.e., the double sigmoid function exhibits linear characteristics in the interval $(t - r_1,\ t - r_2)$. This technique guarantees that the set of normalized scores has a common numerical range [0, 1]. But, it requires careful tuning of the parameters t, r_1, r_2 to obtain good efficiency.

6. Fusion

The choice of the fusion strategy mostly depends on the application scenario of the system. For example it could be preferable to have a high security access to restricted areas, or just to provide a privileged access to a sub-set of users (e.g. fast track in airports). In our experiments we tested two different fusion techniques usually adopted for multimodal biometric systems, i.e. fusion at feature level and fusion at score level.

Fusion at feature level has been performed as suggested by Ross and Govindarajan in [27], in this work the two feature vectors obtained from hand and face biometrics, are first concatenated and then feature selection is applied on the so obtained vector. The selection criterion is the average of Genuine Accept Rate (GAR) corresponding to 4 different FARs (0.05%, 0.1%, 1%, 10%). In our paper, we applied a similar technique, we concatenated the feature vectors obtained from the iris recognition module and the sensor recognition module and used the resulting vector as the new feature vector.

For the fusion at score level, we first computed the distance matrices for the two recognition modules and then we tested different score normalization techniques to be applied before combining the scores.

7. Experimental results

7.1. Data acquisition and preprocessing

We tested our approach on MICHE database [3][2]. It includes more than 3000 eye images, taken from 75 individual subjects. Eye images were acquired with three different mobile devices, namely tablet Samsung Galaxy Tab 2, Apple iPhone5, and Samsung Galaxy S4 smartphone, under uncontrolled acquisition conditions both outdoor and indoor, leading to a very heterogeneous database. In our experiments we employed only the images captured by the two smartphones, with both their frontal and rear cameras (for a total of four different sensors). In Figure 5 some examples of MICHE eye images are presented, the images are rather different from each other due to the technical features characterizing each sensor and due to the uncontrolled acquisition conditions.

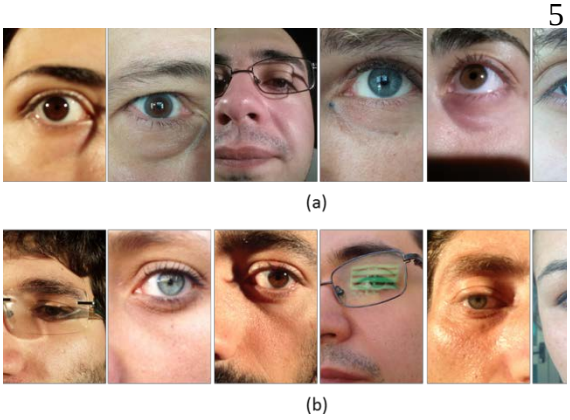

Figure 5 Examples of images in MICHE: (a) captured from Galaxy S4, (b) captured from iPhone5. In both rows odd positions correspond to indoor images and even positions to outdoor ones.

For sensor recognition no preprocessing is needed, the picture as it is, is submitted to the ESPN extractor. For iris recognition some further steps are required, in fact we need to extract the iris from the whole picture that contains also other information, e.g. the periocular area and part of the face, that we do not need in the following steps. As providing an automatic segmentation algorithm is beyond the scope of this paper, we manually segmented the images and we did not remove occlusions due to reflections, eyelids, etc. After iris segmentation we performed a transformation from Cartesian to polar coordinates in order to obtain a rectangular shape of the iris on which is easier to apply the CSUM algorithm.

7.2. Sensor recognition

It is well known [20,21] that device recognition based on SPN extraction is a very robust technique. However we investigated its use on mobile devices, which are limited in terms of memory and computational power. In this section we will present different experiments on sensor recognition in order to show the robustness of this technique even if applied on a small part of the image. In appendix A of [20], it is suggested to process large images by blocks of 512x512 pixel, but during our experiments we observed that using just one block, the same for all images, is sufficient to obtain a RR of 98%, for this reason, in our experiments we extracted from all the images a block of size 512x512 starting from the top-left corner of the photo.

As stated before, with the SPN-based technique it is possible to distinguish which is the device that captured a given photo even among different devices that embed sensors of the same model. In order to test the performance of the sensor recognition algorithm, we extracted the RSPNs of four cameras employed to acquire MICHE images: Galaxy S4 front camera, Galaxy S4 rear camera, iPhone 5 front camera and iPhone 5 rear camera. While collecting the images for MICHE database, the iPhone 5 was changed with another device of the same model, this means that from the subject with ID = 49, photos were acquired with the new device, and since we extracted the RSPN from the new device, pictures relative to IDs less than 49, should be detected as impostors by the system. Moreover, the presence of unrolled IDs in the probe, i.e. pictures captured with a device of which we do not have the RSPN (“old” iPhone 5) in the Gallery, makes the system performance assessment more reliable. We used the RSPNs extracted from the four cameras as Gallery set and the ESPNs extracted from 579 photos selected from the MICHE database as Probe set. The system obtained a RR of 98% and a very low average FAR of about 5%, AUC is equal to 0.99. Results for sensor recognition are shown in Figure 6.

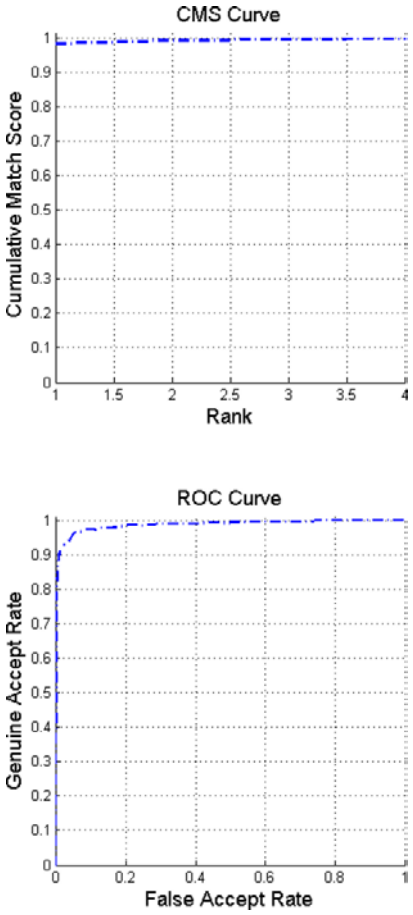

Figure 6 Sensor recognition performances

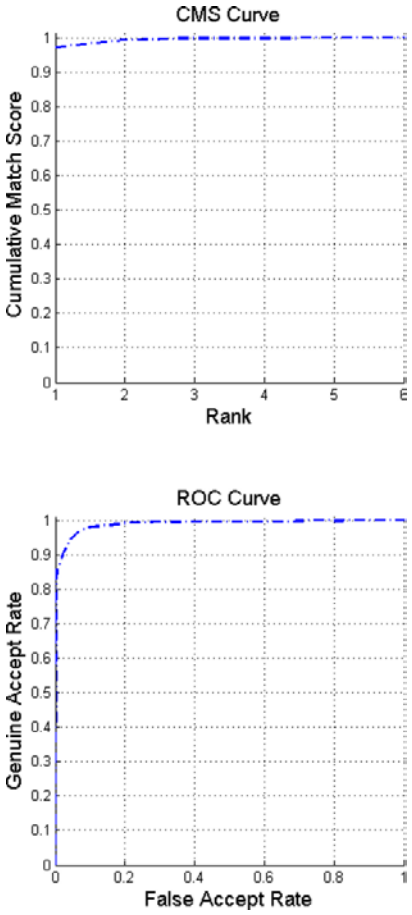

Figure 7 Sensor recognition performance: experiment on same camera model

7.2.1. Different sensors of the same model

In the following experiment, we employed three smartphones of the same model, namely Samsung Galaxy S4. We compared a total of 6 different sensors, as we extracted the RSPNs from both the frontal and the rear cameras of the three devices. Gallery set is thus composed by the 6 RSPNs while the Probe consists in 1297 images from MICHE database, the ones captured with Galaxy S4 smartphone. As just one of the three smartphones was actually employed in MICHE acquisition, the system should correctly assign all the 1297 images to the correct device. Results, see Figure 7, shows that the Recognition Rate (RR) is 97% and the Area under the ROC curve (AUC) is equal to 0.99.

7.2.2. Reference Sensor Pattern Noise Extraction

In order to extract the RSPN of a sensor it is worth employing an high number on images (recommended more than 50) of the blue sky because this kind of pictures do not contain details that, as the sensor noise, are located in the high frequencies of the images and can be confused with it [20]. However, we imagined that for a user, collecting images of the blue sky could be difficult, for this reason and as suggested in [24], we compared the performance of the sensor recognition system when using blue sky images or using any kind of pictures to extract the RSPN of the sensor, results are presented in Figure 8.

Values obtained for RR and AUC are very close, with RR = 98% for both the experiments and AUC = 0.92 and AUC = 0.93 for the case in which RSPN is extracted from blue sky images and the case in which it is extracted from any kind on pictures, respectively. We performed these experiments on 1297 images captured by the Galaxy S4, and compared them with 3 RSPNs from three frontal cameras of three different Galaxy S4.

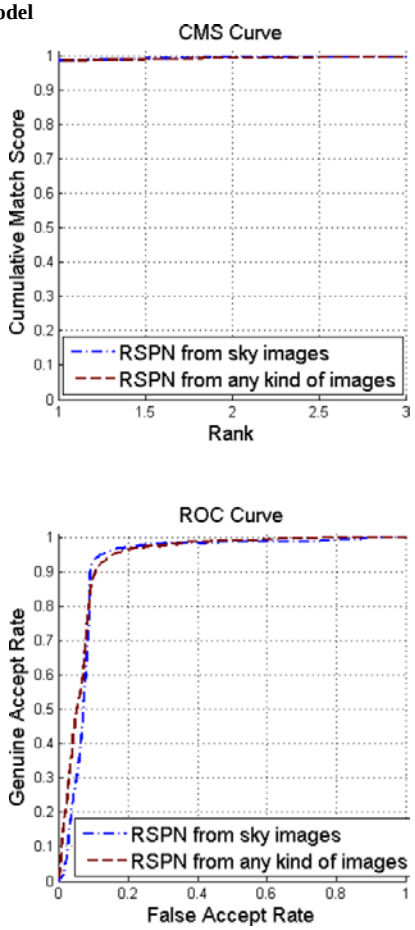

Figure 8 Sensor recognition performances: in red the curves relative to the case in which the RSPN is extracted from blue sky images and in blue

the curves relative to the case in which the RSPN is extracted from any kind of images captured by the sensor.

7.3. Iris recognition

In this section we present the performances of the iris recognition module. Images selected from MICHE dataset were split in Gallery and Probe sets. Probe set is composed by 298 iris images belonging to 75 subjects. The Gallery is composed by 150 iris images, we selected only a part (half) of the 75 individual subjects composing MICHE dataset, we did it in order to simulate the attempt of not enrolled subjects to access the system. Results are shown in Figure 9, with RR = 85% and AUC = 0.77. Performances are poor due in part to the noise introduced by the acquisition in uncontrolled settings (e.g. specular reflections, eyelids and eyelashes occlusions, etc.) and in part to the sensor interoperability problem, in fact MICHE images of the same iris often appear very different because they were acquired by different sensors.

7.4. Fusion at feature level

In this paragraph we present the experiments relative to the combination of iris recognition and sensor recognition. We first tested the fusion at feature level, concatenating the feature vectors extracted from the two recognition modules. To compare the new feature vectors obtained, as the two algorithms employed for iris and sensor recognition use different matching techniques, namely hamming distance for iris recognition and correlation for sensor recognition, we tested both approaches. Results are presented in Figure 10: performances are very close with AUC of about 0.93 for both distance metrics and RR of 23% obtained by Hamming distance and RR of 20% when using Correlation.

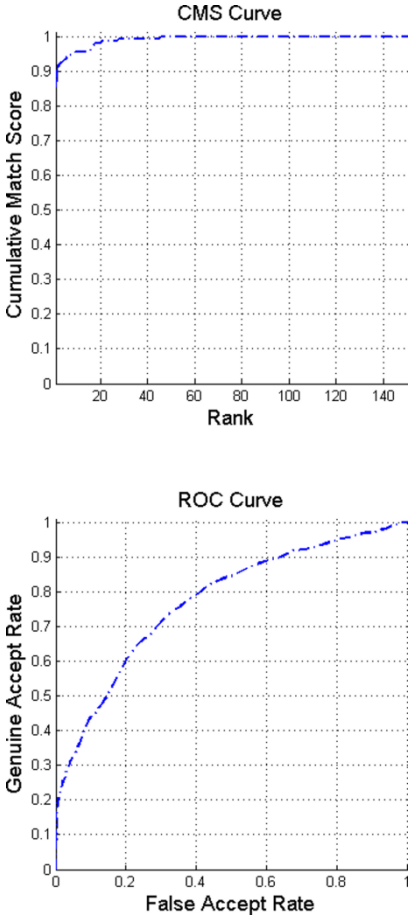

Figure 9 Iris recognition performances affected by sensor interoperability problem

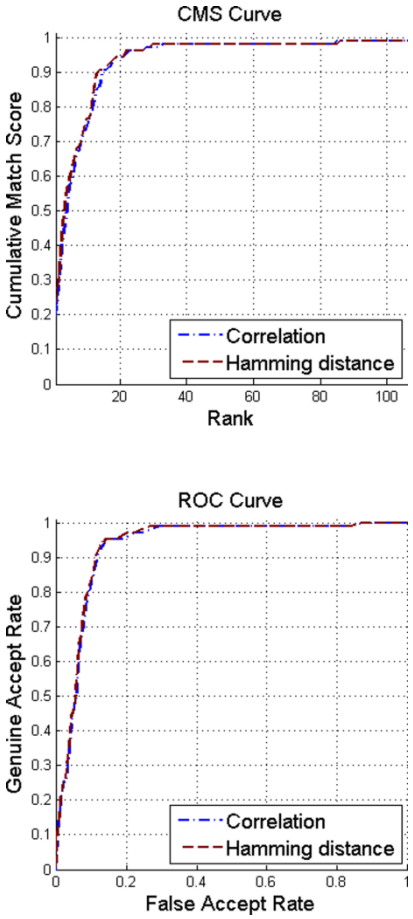

Figure 10 Fusion at feature level performances.

7.5. Fusion at score level

We also tested fusion at score level, we computed the distance matrices of the two recognition modules and then we combined the scores obtained averaging them. Before combining the score, a score normalization step is required. We tested different normalization techniques and we report the results obtained in Figure 11. The best performances are obtained via fusion at score with Max-Min score normalization: the AUC value is equal to 0.98 while the RR is 86%.

The results obtained show that fusion at score level is more suitable than fusion at feature level for this kind of system. Since the system recognize a couple of entities very different in nature, we assigned the same weight both to iris module scores and sensor module ones, in order to avoid the system to be biased towards recognizing the iris or the device.

7.6. Noise response

In this section we want to highlight the advantage in using the sensor recognition module in combination with biometric recognition. It has really high and robust performances as shown in section 7.2. Here we present an example that confirms what stated before: if we submit to the system more challenging pictures, e.g. eye pictures with strong noise due to large specular reflections, important occlusions etc., the iris recognition module performances drop while sensor recognition performances stay the same. We employed pictures from MICHE database acquired outdoor, in Figure 12 we can see that, with respect to performances obtained on indoor photos, the RR drops from 85% to 21% and the AUC from 0.77 to 0.67. Figure 13 shows that even on outdoor pictures, sensor recognition performances remain high, with RR = 98% and AUC = 0.99.

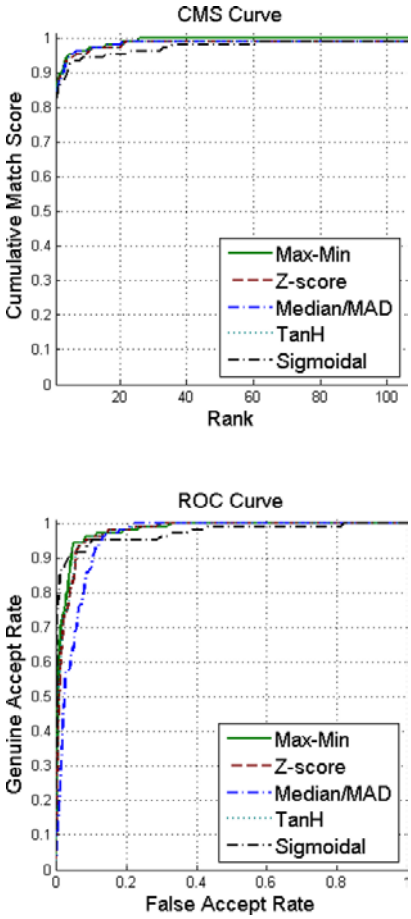

Figure 11 Fusion at score level performances

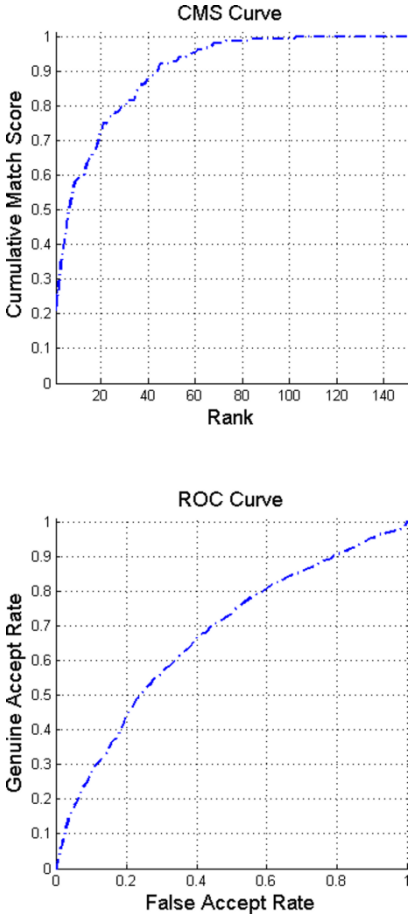

Figure 12 Iris recognition performances on outdoor images.

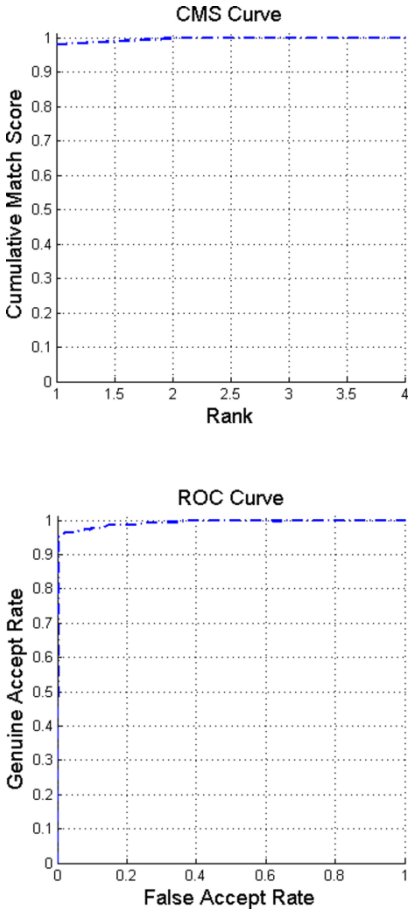

Figure 13 Sensor recognition performances on outdoor images.

8. Conclusions

The main goal of the paper is to present a novel idea, a multimodal system based on the combination of iris recognition and device recognition and demonstrate its applicability on smartphones. The advantages of this technique are manifold: high security level; the genuine sample consists in the couple user-device, making more difficult the spoofing process; good trade-off between accuracy and ease of use; the two recognition processes are applied on a single photo at once; performance of iris recognition and, in particular, of sensor recognition, is very high.

We performed an extensive series of experiments to show that the SPN-based technique can be reliably applied on smartphones. This technique is based on the discrete wavelet transform. Large images, as those captured by nowadays smartphones, should be processed by blocks of 512x512 pixel. E.g. for a picture of MICHE database acquired by Galaxy S4, of size 2322x4128 pixel, the ESPN extraction process should be applied around 36 times. This would require a high computational cost not suitable for the application of this technique on smartphones, that are still limited in terms of memory and computational power. To speed up the process and reduce the computational cost we used just a small block of the image and we obtained a Recognition Rate (RR) of 97% and an Area under the ROC curve (AUC) equal to 0.99.

For what concerns the performances obtained by the fusion of device and iris recognition modules, we want to clear up that the performances were not expected to outperform the single modules. The reason is that the two modules recognize two different entities and their fusion recognize a combination of entities. This is different from a multi-biometric system were two

user's traits are combined to recognize his/her identity (same entity) and thus the performances should outperform the single modules. Indeed the performance were expected to be limited by the weakest module: the iris recognition module. However, we can see that, as reported in Table 1, fusion performances greatly outperform iris recognition ones. The reason is that the CSUM algorithm employed for iris recognition, suffers for the sensor interoperability problem introduced before. Thus, when it is required to recognize the iris despite the sensor that acquired it, its performances drop down. On the contrary, on the fusion scenario, it is required to distinguish between irises acquired with different sensors, getting around the sensor interoperability problem, and obtaining better performances from the iris recognition module.

Table 1 Experimental results summary

	EER	**avg FAR**	**avg FRR**	**RR**	**AUC**
Iris	0.2951	0.2747	0.6044	0.8553	0.7723
Sensor	0.0447	0.0537	0.5592	0.9825	0.9883
Fusion	0.0569	0.2758	0.3590	0.8585	0.9797

Finally we demonstrated that the more the quality of acquired iris degrades the more the SPN is important in a verification process. In fact the experiments presented show that sensor recognition has very high and robust performances.

8.1. Future Implications and Open Issues

This novel system can provide a more secure authentication process without the disadvantage of requiring dedicated sensors. The authentication process is fast and easy, in one single shot the user can get authenticated via his/her iris and his/her smartphone. The smartphones are nowadays strictly related to the owners, and in many companies smartphones are provided to the employees and they are required to bring them during the working hours. This is the perfect scenario in which passwords, tokens or badges can be replaced by the authentication system proposed here. And it is worth noticing that this kind of system is particularly suitable for this scenario because it can distinguish devices of the same model with high accuracy, as shown in paragraph 7.2.1, and it is very likely that the devices provided by a company are of the same model or belong to a restricted set of models.

The use of the recognition of the smartphone in addition to the iris, makes the spoofing attacks more difficult to be carried on. The opponent has to spoof both modules, that even if still possible, it is more complicated than spoofing a system based only on biometric recognition.

Still to be investigated is the possibility of improving system robustness against attacks by adding anti-spoofing techniques, e.g. a liveness detector for the iris recognition module.

References

[1] Authentication systems' security levels: http://www.ffiec.gov/pdf/authentication_guidance.pdf

[2] MICHE database: http://biplab.unisa.it/MICHE/database/

[3] M. De Marsico, M. Nappi, D. Riccio, H. Wechsler, "Mobile Iris Challenge Evaluation - MICHE - I, Biometric iris dataset and protocols", Pattern Recognition Letters, Volume 57, 1 May 2015, pp. 17-23.

[4] B. Chen, J. Shen, H. Sun, "A fast face recognition system on mobile phone", Systems and Informatics (ICSAI), 2012 International Conference on. IEEE, 2012, pp. 1783-1786.

[5] K. Imaizumi, V. G. Moshnyaga, "Network-based face recognition on mobile devices", Consumer Electronics, Berlin (ICCE-Berlin), 2013. ICCEBerlin 2013. IEEE Third In-ternational Conference on, pp. 406-409.

[6] S. Barra, M. De Marsico, C. Galdi, D. Riccio, H. Wechsler, "FAME: Face Authentication for Mobile Encounter", Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2013 IEEE Workshop on, pp. 1-7.

[7] D.H. Cho, K.R. Park, D.W. Rhee, Y.G. Kim, J.H. Yang, "Pupil and iris localization for iris recognition in mobile phones", Proceedings of the SNPD (2006), pp. 197–201

[8] D.H. Cho, K.R. Park, D.W. Rhee, "Real-time iris localization for iris recognition in cellular phone", International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (2005), pp. 254–259.

[9] Samsung: http://www.samsung.com/

[10] CASIA database: http://www.sinobiometrics.com

[11] M. Frucci, C. Galdi, M. Nappi, D. Riccio, G. Sanniti di Baja, "IDEM: Iris DEtection on Mobile devices", 22nd International Conference on Pattern Recognition, ICPR2014, August 24-28, 2014, pp. 1752-1757

[12] A. F. Abate, M. Frucci, C. Galdi, D. Riccio, "BIRD: watershed Based IRis Detection for mobile devices", Pattern Recognition Letters, Volume 57, 1 May 2015, pp. 43–51

[13] D.S. Jeong, H.A. Park, K.R. Park, J. Kim, "Iris recognition in mobile phone based on adaptive Gabor filter", International Conference on Advances on Biometrics (ICB '06)3832LNCS (2006), pp. 457–463.

[14] K.R. Park, H. Park, B.Y. Kang, E.C. Lee, D.S. Jeong, "A study on iris localization and recognition on mobile phone", Eur. J. Adv. Signal Process. (2007), pp. 1–12.

[15] J.S. Kang, "Mobile iris recognition systems: an emerging biometric technology", Interna-tional Conference on Computational Science (ICCS) (2010), Procedia Computer Science, Volume 1, Issue 1, May 2010, pp. 475-484.

[16] S. Barra, A. Casanova, F. Narducci, S. Ricciardi, "Ubiquitous iris recognition by means of mobile devices", Pattern Recognition Letters, Volume 57, 1 May 2015, pp. 66-73.

[17] A. F. Abate, M. Nappi, F. Narducci, S. Ricciardi, "Fast Iris Recognition on Smartphone by means of Spatial Histograms", Biometric Authentication, Lecture Notes in Computer Science, Springer International Publishing, 2014, pp. 66-74.

[18] M. De Marsico, C. Galdi, M. Nappi, D. Riccio, "FIRME: face and iris recognition for mobile engagement", Image Vis. Comput. (2014) Volume 32, Issue 12, December 2014, pp. 1161–1172.

[19] C.-T. Li, "Source camera identification using enhanced sensor pattern noise", IEEE Trans-actions on Information Forensics and Security 5(2): pp. 280-287, 2010.

[20] J. Lukás, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise", IEEE Trans. Inf. Forensics Security, vol. 1, no. 2, pp. 205–214, Jun. 2006.

[21] M. Chen, J. Fridrich, M. Goljan, J. Lukás, "Determining image origin and integrity us-ing sensor noise", IEEE Trans. Inf. Forensics Security, vol. 3, no. 1, Mar. 2008, pp. 74–90.

[22] J.-G. Ko, Y.-H. Gil, J.-H. Yoo, K.-I. Chung, "A novel and efficient feature extraction method for iris recognition," ETRI J. 29 (3), 2007, pp. 399–401.

[23] W. Taktak, J.-L. Dugelay, "Digital Image Forensics: A Two-Step Approach for Identifying Source and Detecting Forgeries", The Era of Interactive Media, Springer New York, 2013, pp. 37-51.

[24] J. A. Redi, W. Taktak, and J.-L. Dugelay, "Digital image forensics: a booklet for beginners", Multimedia Tools and Applications, vol. 51, no. 1, 2011, pp. 133–162.

[25] H. Proença, L. A. Alexandre, "The NICE.I: Noisy Iris Challenge Evaluation – Part I", IEEE First International Conference on Biometrics: Theory, Applications and Systems, BTAS 2007, pp. 1-4.

[26] iPhone 6 Touch ID: https://www.apple.com/iphone-6/touch-id/

[27] A. Ross, R. Govindarajan, "Feature Level Fusion Using Hand and Face Biometrics", SPIE Conference on Biometric Technology for Human Identification II, (Orlando, USA), March 2005, Vol. 5779, pp. 196-204.

[28] A. Ross, A. Jain, "Biometric Sensor Interoperability: A Case Study In Fingerprints", International ECCV Workshop on Biometric Authentication (BioAW), (Prague, Czech Republic), Springer Publishers, May 2004, LNCS Vol. 3087, pp. 134-145.

[29] A. Jain, K. Nandakumar, A. Ross, "Score normalization in multimodal biometric systems", Pattern Recognition, Volume 38, Issue 12, December 2005, pp. 2270-2285.

[30] F.R. Hampel, P.J. Rousseeuw, E.M. Ronchetti, W.A. Stahel, “Robust Statistics: The Approach Based on Influence Functions”, Wiley, New York (1986)

[31] R. Cappelli, D. Maio, D. Maltoni, “Combining fingerprint classifiers”, Proceedings of First International Workshop on Multiple Classifier Systems (2000), pp. 351–361