Iris Liveness Detection for Next Generation Smartphones

Shejin Thavalengal, *Student Member*, IEEE, Tudor Nedelcu, *Student Member*, IEEE,
Petronel Bigioi, *Senior Member*, IEEE, and Peter Corcoran, *Fellow*, IEEE

Abstract — *This paper presents a novel liveness detection method that exploits the acquisition workflow for iris biometrics on smartphones using a hybrid visible (RGB)/near infra-red (NIR) sensor. These devices are able to capture both RGB and NIR images of the eye and iris region in synchronization. This multi-spectral information is mapped into a discrete feature space. An intermediate classifier which uses a distance metric close to Jenson-Shannon divergence is employed to classify the incoming image. Further, a fast, multi-frame pupil localization technique using one-dimensional processing of the eye region is proposed and evaluated. This is used to analyze the pupil characteristics of the images classified as 'live' in the previous stage. It is shown that such an analysis could detect presentation attacks, even with a 3-D face model made of materials that has properties similar to human skin and the ocular region*[1].

Index Terms — Smartphone, consumer biometrics, iris recognition, liveness.

I. Introduction

A. Smartphones

Since the introduction of first smartphone in 1994, there has been a rapid evolution of smartphone technology to a point where it plays a central role in our day to day life [1]. Approximately 2 billion people will be using a personal smartphone in 2016, which is expected to grow to a third of the world's population in 2018 [2]. These devices have become much more than a computer, providing the functions of a phone, a personal database, an infinite jukebox, a camera, a hub for location based services and a gateway to all the information in the world [3]. It is speculated that the smartphone's role as a constant companion, helper, coach and guardian has only just begun [4].

[1] This work is supported by the Irish Research Council's Employment based PhD program and part funded by FotoNation Ireland.

Shejin Thavalengal is with the FotoNation Ltd, Galway, Ireland and CoEI, National University of Ireland Galway, University Road, Galway, Ireland (e-mail: sthavalengal @fotonation.com).

Tudor Nedelcu is with the FotoNation Ltd, Galway, Ireland and CoEI, National University of Ireland Galway, University Road, Galway, Ireland (e-mail: v-tnedelcu@fotonation.com).

Petronel Bigioi is with the FotoNation Ltd, Galway, Ireland and CoEI, National University of Ireland Galway, University Road, Galway, Ireland (e-mail: pbigioi@fotonation.com).

Peter Corcoran is with the college of Engineering and Informatics, National University of Ireland, University Road, Galway, Ireland (e-mail: peter.corcoran@nuigalway.ie).

The majority of these devices are connected to the Internet all the time. As many as 57% of U.S. smartphone users are reported to carry out online banking through these device [5]. As today's smartphones are increasingly used to transmit sensitive financial and personal information, a reliable assessment of smartphone user's identity is emerging as an important new service. PIN or passwords may not be sufficient for this purpose, but personal biometrics could be effectively used [6].

B. Biometrics

Traditionally biometrics has been used for many years by law enforcement to establish the identity of criminals [7] and by government and industry to secure and restrict access to resources and facilities. In such applications biometric technology is employed in a supervised use-case – another person oversees the authentication procedure in order to ensure correctness. More recently, biometric technology has been employed in non-consensual use-cases such as airports, train-stations and similar public areas. In these example use-cases there is limited scope for 'spoofing' a user biometric.

However when biometric technology is adapted for use on a consumer device, the acquisition process differs significantly as it represents an "unsupervised" use-case. As fingerprint authentication was the first technology to be widely adopted in mobile devices there are many examples of 'spoofing' techniques and countermeasures in the literature [8]–[11].

C. Iris Biometrics on Smartphones

Iris biometrics is widely recognized as being one of the most practical biometrics in use today [12]. The iris of the human eye is the annular region between the pupil and sclera. The iris pattern consists of complex and distinctive ligaments, furrows, ridges, rings, corona, freckles and collarette [13]. Also, the iris is relatively stable over the lifetime of a person starting from the eighth month of gestation and demonstrates high pattern variability, even for identical twins and between the left and right eye of the same person [13]. These characteristics make the iris a very suitable candidate for biometric user authentication in smartphones.

The implementation of iris biometrics on smartphone devices has recently become an emerging research topic [14]–[16]. As the use of iris biometrics on smartphone devices becomes more widely adopted, it is to be expected that there will be similar efforts in the research community to beat the biometric by exploring new spoofing methods and this will drive a corresponding requirement for new liveness detection methods.

II. Iris Liveness Detection: A Brief Overview

Smartphone user authentication using iris biometrics is a remote and unsupervised form of authentication. In other words, only the person performing the authentication needs to be present during the process workflow. As a consequence, it is more susceptible to spoofing of the biometric input than traditional authentication techniques such as PIN entry at a point-of-sale terminal where the sales clerk is present.

Spoofing attack on biometric system is an artificial mimic of a real biometric to gain access to the device and its services. This become worrisome as the iris biometric sample can be recorded without user co-operation. Hence it is essential to build in protection against attacks into such a system. Various types of spoofing include presenting a picture, a recorded video or a high quality iris image kept in front of original eye while trying to use iris authentication. These attacks are collectively called 'presentation attacks' [17].

Liveness detection is an anti-spoofing technique to determine if the biometric being acquired is an actual measurement from a live person who is present at the time of capture [18]. Arguably, human supervision can be the most effective way for detecting such presentation attacks and widely used in many applications including UAE border control program. But, it is impractical in the case of smartphones and other consumer electronic devices. Hence, effective automatic liveness detection is necessary.

Czajka [17] categorizes the automatic liveness detection techniques into three categories: (a) extraction of intrinsic properties of a living body, (b) analysis of involuntary signals and (c) challenge-response method. The extraction of intrinsic properties includes analyzing spectrographic properties of the human eye, analyzing red-eye effect or analyzing 3-D curvature of iris surface. Examples for analyzing involuntary body signals include eyelid movements and hippus. The third category mainly considers user's response when prompted to carry out some tasks like blinking, or looking at a different direction. A detailed literature review of iris liveness detection can be found in [17].

Even though liveness detection is an essential part of iris recognition system as a countermeasure against spoofing, it comes with the cost of an increase in processing time, increase in hardware or software and negative effect on recognition performance [18].

III. Liveness Detection on Smartphones

Iris spoof detection is rapidly growing recently [17]–[20]. All of the existing work in the literature can be further classified into two classes – (i) techniques which require special hardware or user interaction [17] and (ii) algorithms designed to work on static images/videos (such as high quality printed iris image, iris images presented on a screen, or a video stream) [20]. Techniques which need special hardware or user cooperation may not be an ideal solution on smartphones as these will increase the cost and decrease the usability.

Also, if the techniques demand considerable amount of user effort to make a decision, they will not be adopted for everyday use by the consumers. Note that, this is still an emerging technology and even if iris authentication is feasible there are still ease-of-use issues to overcome in order to achieve a broad adoption of the technology in consumer market.

Hence, simple, cost-effective yet powerful liveness detection should be incorporated on smartphones. Such a technique ideally should not require any additional hardware or user interaction and should be computationally light enough to be embedded in the smartphone camera pipeline or dedicated iris recognition digital signal processors.

Current research on iris liveness techniques is limited to either NIR iris recognition systems or the visible spectrum ones. But, smartphones powered with NIR iris recognition is already in market [14]. These devices use one dedicated camera for iris recognition. But in the near future, it is speculated that these smartphones will be employing one hybrid front facing camera for iris recognition as well as general purpose front camera use such as video call and selfie imaging [16], [21]. Example for such a device is shown in in Fig. 1.

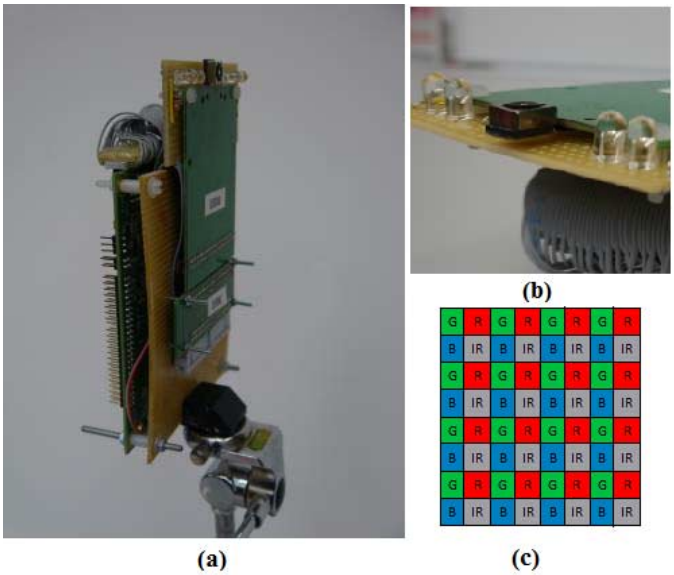

Fig. 1. RGB/NIR hybrid camera for smartphones. (a) Prototype device. (b) Smartphone form factor optics. (c) Hybrid sensor used in the prototype device - R, G, B and IR represents Red, Green, Blue and Infra-Red sampling filters respectively.

Hence, harnessing the existence of these two modalities for iris liveness detection may provide efficient liveness detection capabilities to such devices without the need of any additional hardware. The feasibility of this approach is examined in this article.

The specific contributions of this manuscript are listed as follows: (i) novel iris liveness detection for next generation smartphones is presented. The proposed liveness detection technique harnesses the workflow of hybrid RGB/NIR sensor for smartphone iris recognition. Such a technique provides

efficient liveness detection capabilities without the need of any additional hardware. (ii) Computationally efficient iris localization technique presented previously [22] is adapted for pupil analysis for liveness detection. Efficacy of the proposed liveness detection technique is tested over various attack scenarios including the attack with a mannequin, which has engineered artificial eyes to duplicate the optical behavior of human eyes.

IV. Proposed Liveness Detection Process

A two stage iris liveness detection technique is proposed in this section. The proposed technique, depicted in Fig. 2, is motivated by the novel RGB/NIR hybrid sensor and associated workflow presented in the previous work [16], [21].

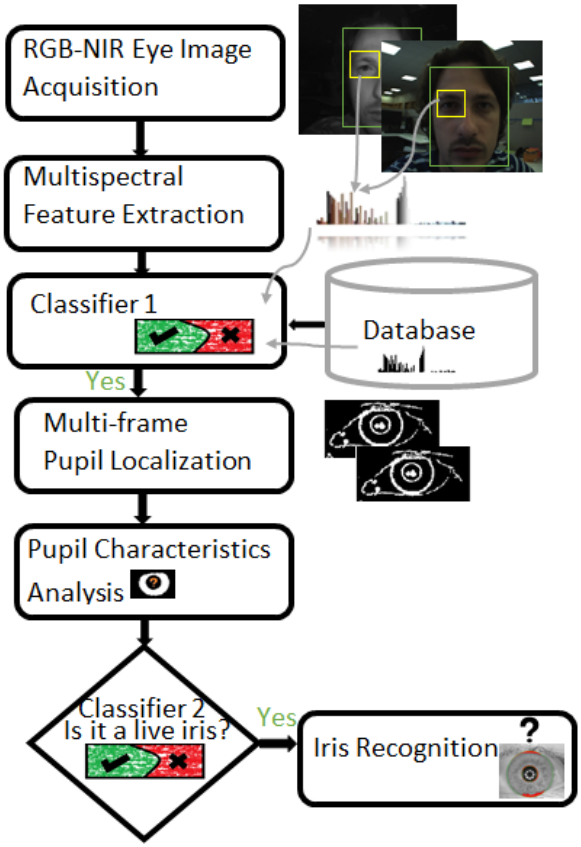

Fig. 2. Workflow of the proposed liveness detection technique.

The various steps of the proposed technique are described in the following subsections.

A. RGB-NIR Image Acquisition

Once the user has pressed the 'wake up' or 'unlock' key of the smartphone, the device will start acquiring eye images for further processing. State of the art face detection, eye detection and eye tracking techniques on the image stream in the visible wavelength are employed to obtain a sequence of good quality in-focus eye regions [23]. Once a good quality eye region is detected, both RGB and NIR images of the eye region are acquired. The RGB and NIR image formation can be explained as [24],

$$I_v = \int_{\lambda_v}\int_{p} E(\mathrm{p},\lambda_v)R(p)Q(\lambda_v)\,\mathrm{dpd}\,\lambda_v \quad (1)$$

$$I_i = \int_{\lambda_i}\int_{p} E(\mathrm{p},\lambda_i)R(p)Q(\lambda_i)\,\mathrm{dpd}\,\lambda_i \quad (2)$$

Where $I_v \in \mathbb{R}^{m\times n}$ is the RGB image, $I_i \in \mathbb{R}^{k\times l}$ is the NIR image, $\lambda_v \in [350\,\mathrm{nm}, 700\,\mathrm{nm}]$, $\lambda_i \in [750\,\mathrm{nm}, 900\,\mathrm{nm}]$ are the wavelength range of RGB and NIR image respectively. Also, p is the spatial domain of the sensor, R is the spatial response of the sensor, E is the irradiance and Q is the quantum efficiency of the device. $I_i \in \mathbb{R}^{k\times l}$ is demosaiced/interpolated to obtain $m = k, n = l$,

These two images, I_v and I_i are fused to make a hyperspectral image I_h

$$I_h = \Gamma(I_v, I_i) \quad , \quad (3)$$

Where $I_h \in \mathbb{R}^{m\times n\times 4}$ and Γ is the fusing operator. Broadly speaking I_h represents the ambient light and surface reflectance on the eye at four different wave bands (Blue, Green, Red and NIR). This image is further processed to minimize the effect of ambient light by the metadata obtained from the camera, which is used for white balancing and auto exposure of the RGB camera image stream [25].

B. Multispectral Feature Extraction

It can be observed from the previous discussion that the hyperspectral image I_h is made of four image planes $(I_{c1}, I_{c2}, I_{c3}, I_{c4})$ of size $m\times n$ representing four different wavebands (Blue, Green, Red and NIR). The pixels in each channel are clustered separately to α predefined clusters as,

$$I^u_{cj} = \Omega(I_{cj}), \quad (4)$$

where $I^u_{cj} \in [1,\alpha]^{m\times n}$ represents the label of the cluster corresponds to the pixels in $I_{cj}, j \in [1,4]$ denotes the image channel and Ω is the clustering operator, In this work, a nearest neighborhood clustering operation which groups the pixels in each plane into one of the α cluster based on the intensity value is used as Ω. Considering the dimensionality and computational complexity, $\alpha = 8$ is chosen. The label maps are further concatenated as,

$$I^u_h = \Gamma'(I^u_{c1}, I^u_{c2}, I^u_{c3}, I^u_{c4}), \quad (5)$$

where Γ' is the concatenation operator. Due to different combinations of clustering made by the concatenation of channels, each element in I^u_h can have one of the $s = \alpha^4$ unique combinations. The normalized frequency distribution of each combination is calculated using the transform operator H

$$H :\rightarrow I^u_h = F, \quad (6)$$

where $F = (\mathrm{f}_1, \mathrm{f}_2, .., \mathrm{f}_s)$ is the number of times each unique

cluster combination appeared in I_h^u. This mapping is used as the feature vector for further processing.

This specific feature extraction technique is chosen as it represents the unique distribution of information across various image planes in the hyperspectral image I_h. Also, the proposed technique is found to be computationally inexpensive and generate a compact feature vector of 4096 elements.

C. Intermediate Decision Making

The training feature vectors for the live user are calculated while the iris enrolling stage. A one class classifier is trained on this live feature vector. Such a classifier will be able to predict whether the incoming images are most likely belonging to live person or a presentation attack.

The deviation (error) d when the feature vectors of the query image F^q from the distribution of the representative feature vectors in the trained model F^{db} is computed from a Bayesian point of view. That is, the distance is measured as the square root of the entropy approximation to the logarithm of evidence ratio when testing whether the query image can be represented as the same underlying distribution of the live images [26]. This can be mathematically represented as,

$$d^{q,db} = \sqrt{D\left(F^q \,\|\, \frac{1}{2}\left(F^q + F^{db}\right)\right) + D\left(F^{db} \,\|\, \frac{1}{2}\left(F^q + F^{db}\right)\right)} \quad (7)$$

$$d^{q,db} = \sqrt{\sum_{z=1}^{s}\left(f_z^q \log\frac{2f_z^q}{f_z^q + f_z^{db}} + f_z^{db} \log\frac{2f_z^{db}}{f_z^q + f_z^{db}}\right)} \quad (8)$$

Here, $D\left(F^q \,\|\, F^{db}\right)$ is the Kullback-Leibler divergence of F^{db} from F^q, which is a measure of information lost when the database feature vector F^{db} is approximated from the query feature vector F^q. The above presented choice of distance metric $d^{q,db}$ is based on the observations that it is a close relative to Jenson - Shannon divergence and an asymptotic approximation of χ^2 distance. Also $d^{q,db}$ is symmetric and fulfills the triangle inequality [26]. If $d^{q,db} < \beta$ - where $\beta \in \mathbb{R}$ is a predetermined threshold- the query image is classified as a live person.

1) Intermediate Decision Making –Experimental Results

Since there are no publicly available databases with RGB-NIR iris image pairs acquired simultaneously and the different presentation attack scenarios in such set up, a database is gathered internally for the proof-of-concept study. This initial database consists of live and presentation attack images of 25 subjects acquired using the device presented in [16], [21].

The RGB and NIR image pairs are acquired at normal office situation with active illumination of 850nm. Examples for the image pairs acquired at different stand-off distance are shown in Fig. 3.

One third of the images in the database are used as training set. The rest of the database is used for testing. Various presentation attacks presented in the literature [20] such as (i) high quality RGB colour prints (in matte paper, and glossy paper), (ii) high quality NIR colour prints, (iii) high quality RGB images presented in a screen and (iv) high quality NIR images presented on a screen are tested on the proposed system.

Fig. 3. RGB-NIR image pairs in the database: (a) RGB image (b) NIR image. Both (a) and (b) are captured synchronously.

The performance of this stage is assessed by ISO/IEC guidelines for different presentation attacks [27]. The 'Normal Presentation Classification Error Rate' (NPCER) which is defined as the proportion of live users incorrectly classified as presentation attack and the 'Attack Presentation Classification Error Rate' (APCER), which is defined as the proportion of presentation attack attempts incorrectly classified as live users. The performance of the overall presentation attack detection system is presented in terms of 'Average Classification Error Rate' (ACER) which is given by the mean of the NPCER and the APCER error rates. The results for various attack scenarios are tabulated in TABLE I.

TABLE I
AVERAGE CLASSIFICATION ERROR RATE (ACER) IN % OF THE PROPOSED CLASSIFIER

Attack Scenario	ACER (%)
Printed RGB Image	1.11
Printed NIR Image	0.37
RGB Image on Screen	0
NIR Image on Screen	0

2) Intermediate Decision Making –Discussion

It can be noted that the intermediate decision making depends on the surface reflection and refraction of the material present in front of the device. Human skin and the ocular region have distinct reflection and refraction properties. This property should be ideally enough to differentiate a presentation attack from the original live faces. One of the most common presentation attacks is presenting a high quality printed photograph or a video stream on a screen in front of the smartphone [20].

As the photographic material (reflective paper or matte paper) and the displays of devices are significantly different from the human skin, the proposed method detects such attacks. This can be observed from Fig. 4, which shows the difference in surface reflection of the above mentioned attack scenarios as compared to a live person. This difference contributes to the high performance of the classifier tabulated in TABLE I.

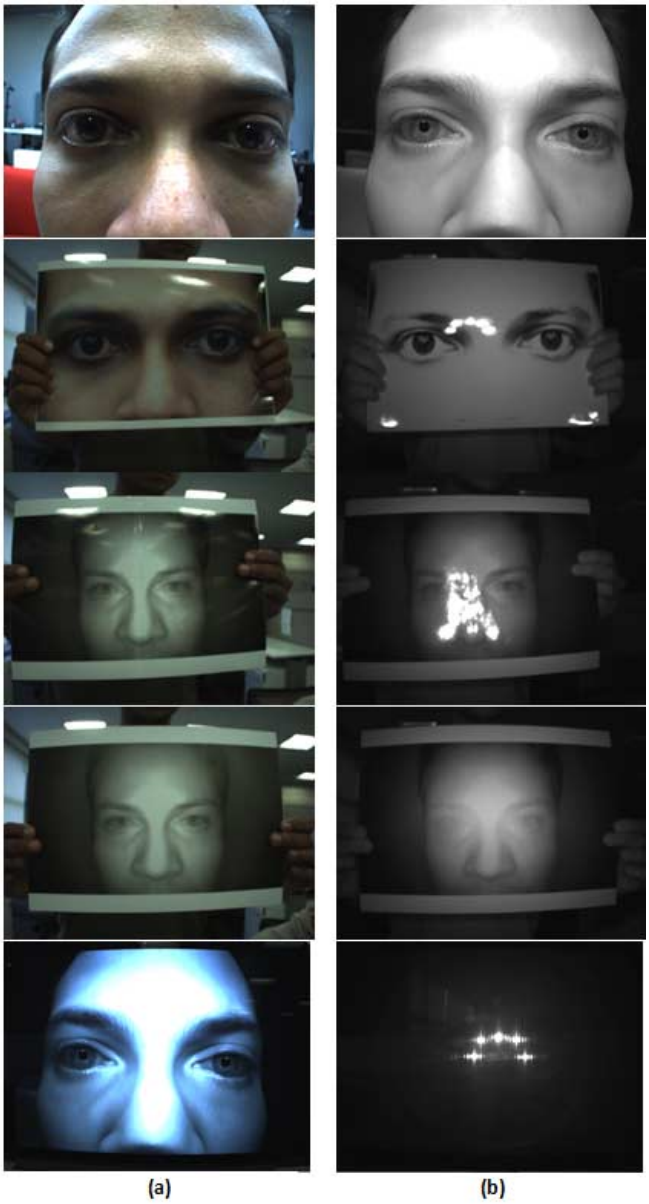

Fig. 4. Different Test Scenarios on the proposed system: (a) Visible Image, obtained by the system and (b) NIR Image acquired. First row is a live person. Second row is attack with a high quality visible printed image. Third row and fourth row are attacks with high quality NIR printed image (on two different photographic papers - glossy and matte). Fifth row is attack with NIR image shown in a laptop screen with high resolution retina display.

One has to note that the spoofing techniques are evolving swiftly and it cannot be guaranteed that spoofing attack using only the known materials will be happening. In order to ascertain this argument, a realistic 3-D face model is used to test the presentation attack detection technique. This mannequin used for test has engineered artificial eyes to duplicate the optical behavior of human eyes such as generating red-eye effect. Also, it has similar material properties of human skin, hair and eye iris region. Such an attack scenario is shown in Fig. 5.

Fig. 5. An example for a sophisticated attack scenario: (a) Visible Image, obtained by the system and (b) NIR Image acquired. First row shows the realistic 3-D face model is used to test the presentation attack detection techniques. Second row is a close up look to indicate the human like skin, hair and ocular properties of the mannequin. Side view is shown in the third row.

From Fig. 5, it can be observed that the mannequin pose a challenging attack scenario as it has human like reflectance and refraction properties. Also, the eye region of the mannequin is capable of producing 'red-eye' like effect (Fig. 5, second row). When the attack scenario with the mannequin was introduced, the APCER is increased to 5.56% while NPCER remained the same. This is due to the fact that this mannequin possesses material properties of human skin and eyes. Hence, mannequin was often misclassified as a live person. Also, it has to be noted that, even though this mannequin was wrongly classified as a live user, it is failed to get authenticated to the system in the iris recognition step. The mannequin presented here can potentially keep a printed contact lens with pattern of the genuine user to get access to the system. Hence, analyzing the spectral response alone will not be sufficient in such sophisticated attack scenarios.

As a consequence, pupil characteristics analysis is recommended as an added security. The pupil analysis is

recommended based on its effectiveness for iris liveness detection presented in the state of the art study by Czajka [17]. Mimicking both the pupil dynamics and material properties of the human skin and ocular region is not feasible with current technologies.

D. Multi-Frame Pupil Localization

Current smartphones have capabilities to acquire 120-240 frames per second, but likely to double with next-generation technology. Hence, it will be practical to capture as many as 30-40 images within the same time window used today to acquire two images. If it is considered 30 frames on average, this will practically give us around 60 images in this time frame (30 RGB and 30 NIR images acquired in synchronization). Hence fast and sufficiently accurate pupil localization is required in this scenario for pupil characteristics analysis. Such a solution is presented by Thavalengal, Bigioi and Corcoran [22].

1) One Dimensional Image Processing for Pupil Localization

The iris region of the eye images coming into the multi-frame pupil localization module may be affected by illumination variations and shadows created by eyelashes. This issue can be addressed using a representation that is less sensitive to illumination variations. One dimensional image processing can be used for this purpose [28]. As compared to other edge based techniques, this technique does not require any thresholding and reduces the smearing of the edges. Further, the choice of this process is motivated by its success in various applications including face recognition and image super resolution [28], [29].

In the one dimensional processing of a given image, a smoothing operator is applied along one direction, and a derivative operator is applied along the orthogonal direction [28]. Let $I \in \mathbb{R}^{m \times n}$ be the cropped eye image (The processing is carried out on the NIR image I_i). The smoothed eye image can be obtained by,

$$I_\theta^s = I\left(\mathrm{x}, \frac{r + x\sin(\theta)}{\cos(\theta)}\right) \otimes \mathrm{S}_\theta(x), \tag{9}$$

where $I_\theta^s \in \mathbb{R}^{m \times n}$ is the smoothed iris image, $S_\theta(\mathrm{x}) \in \mathbb{R}^{m \times 1}$ is the one dimensional smoothing function along a line which has a perpendicular distance of $r \in \mathbb{N}$ from the origin and makes an angle $\theta \in \mathbb{N}$ with the x-axis, $\otimes$ is the one dimensional convolution operator. This convolution operation is carried out for each value of r to obtain the smoothed image I_θ^s. The smoothing function used here is defined by,

$$\mathrm{S}_\theta(x) = \frac{1}{\sqrt{2\pi\sigma_s^{\,2}}} \mathrm{e}^{\frac{-x^2 \sec^2(\theta)}{2\sigma_s^{\,2}}}. \tag{10}$$

Where $\sigma_s \in \mathbb{R}$ is the standard deviation of the Gaussian function used in the smoothing process. A one dimensional derivative operator along the orthogonal direction $\theta + 90^o$ is applied to the smoothed image for different values of r to obtain an intermediate edge gradient image,

$$I_\theta^g = I_\theta^s\left(\mathrm{x}, \frac{r + x\sin(\theta + 90)}{\cos(\theta + 90)}\right) \otimes G_{\theta+90}(x), \tag{11}$$

where ,

$$G_\theta(x) = \frac{x\sec^2(\theta)}{\sqrt{2\pi\sigma_g^{\,6}}} \mathrm{e}^{\frac{-x^2 \sec^2(\theta)}{2\sigma_g^{\,2}}}. \tag{12}$$

Here, $\sigma_g \in \mathbb{R}$ is the standard deviation of the derivative operator. The magnitude representation of edge gradient can be obtained by

$$I_\theta^M = \sqrt{\left(I_\theta^g\right)^2 + \left(I_{\theta+90}^g\right)^2}. \tag{13}$$

A transform operator T is applied on I_θ^M,

$$I_d = T_\delta I_\theta^M, \tag{14}$$

where I_d is the transformed image. The transformation operator T is chosen in such a way that it binarizes the image I_θ^M followed by the detection of largest connected region in the image. $\delta \in \mathbb{N}$ is a threshold in such a way that $n_{min}^p \le \delta \le n_{max}^p$, where n_{min}^p and n_{max}^p are the minimum and maximum number of pixels which could possibly be in the pupil region in this particular frame. From the face and eye tracking meta-data and the camera parameters, an approximate number of pixels in the pupil region can be obtained and the value of δ can be learned for each individual frame [22].

The whole process is depicted in Fig. 6. Note that $\theta = 90^o$ is used in the experiments. Fig. 6(a) represents the original image, Fig. 6(b) and Fig. 6(c) represent the output of one dimensional image processing for the specific angular direction θ and its orthogonal value. Fig. 6(d) is the magnitude image obtained from the result of the one dimensional image processing. The localized pupil is shown in Fig. 6(e).

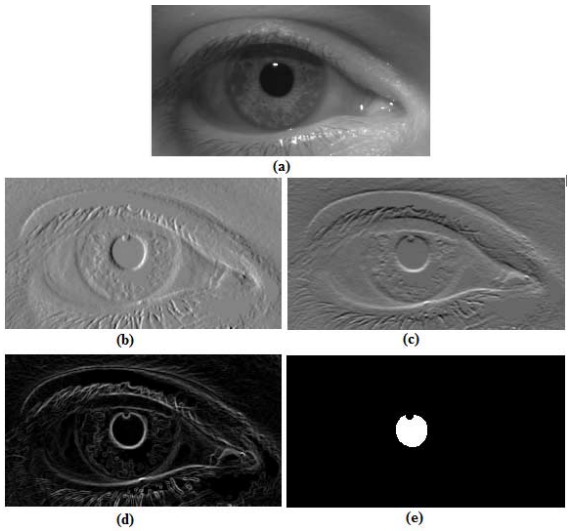

Fig. 6. Pupil Localization Process: (a) Original NIR Image, (b) edge gradient image along one direction, (c) edge gradient image along orthogonal direction, (d) magnitude image and (e) localized pupil.

E. Pupil Analysis

Once the pupil is localized in a sequence of frames, the next step is analyzing the pupil characteristics. Two parameters are used for pupil analysis in this work - (i) pupil area and (ii) pixel intensity in the pupil region. The analysis of pupil area is to note down the eye saccades, hippus and pupil dilation/constriction which may arise naturally as the person move close to the camera.

Also, over a sequence of frames, eye-blinking may happen, which will alter the pupil area. This can be seen from Fig. 7. Fig. 7 represents a sequence on NIR eye images and the pupil localization result of these images. It can be noted from Fig. 7(b) that the area of pupil varies with time which could be measured with sufficient accuracy using the presented localization technique.

The pixel intensity in pupil region will be able to detect the Purkinje image (The virtual image formed by the light reflected from the four optical surfaces of the human eye is called as Purkinje images. Purkinje images are used for various applications including iris liveness detection and eye tracking [30], [31]) or red-eye like effects [32].

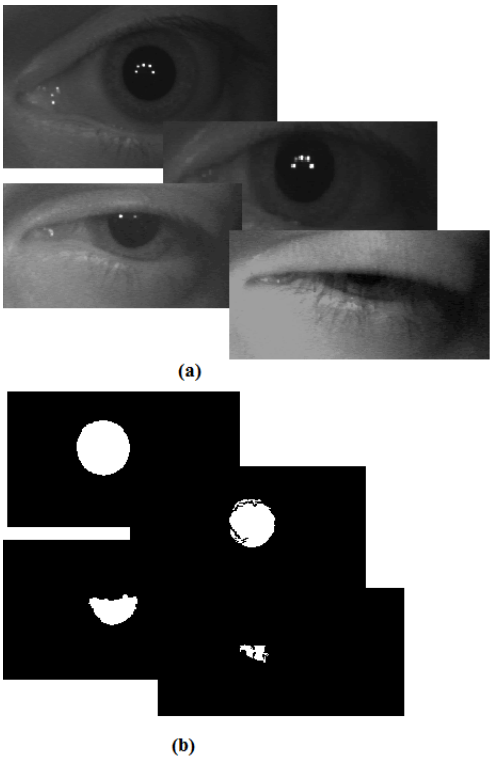

Fig. 7. Pupil Analysis: (a) NIR iris image frames and (b) localized pupil.

F. Decision Making

A binary decision tree is used to classify the image as either a live person or a presentation attack. Binary decision tree is a natural choice when different models (in our case intermediate decision making strategy presented in Section IV *C* and pupil analyses presented in Section IV *E*) become responsible for prediction in different regions of input space [33]. The binary decision tree used one root node and one intermediate node as shown in Fig. 8.

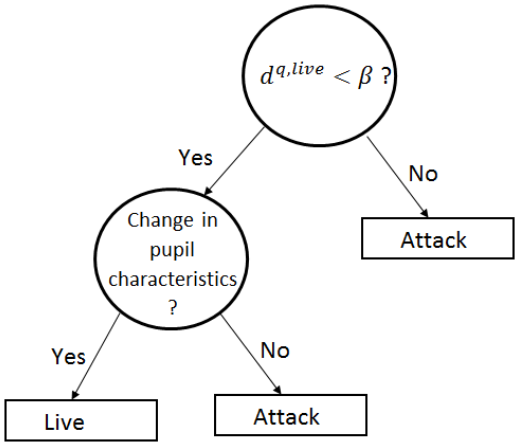

Fig. 8. Binary decision tree used for classifying the input image: the root node measures whether the incoming image is close to a live iris or a presentation attack. The intermediate node uses the pupil analysis information to further classify the output of the first node.

If the resultant image is classified as a 'live' image, the image is passed to the recognition module for iris recognition. The pupil localization result in Section IV *D* is fed to the Iris Recognition unit along with the input images for recognition.

The final performance of this system for all the attack types presented in Fig. 4 and Fig. 5 (including the mannequin attack) is tabulated in TABLE II. These results are compared with the results obtained when pupil analysis was not carried out. From TABLE II, it can be observed that the proposed liveness detection technique, along with the pupil analysis results in accurate iris liveness detection which can be used for next generation smartphone biometrics.

TABLE II
COMPARISON OF ACER (INCLUDING MANNEQUIN ATTACK)

Method	ACER (%)
Without Pupil Analysis	6.3
With Pupil Analysis	0

V. CONCLUSIONS

This article presented a novel liveness detection technique to be used with iris recognition on smartphones. This technique relies on the capability of iris biometrics enabled smartphones to acquiring RGB and NIR image pairs simultaneously. It has been demonstrated that harnessing the capabilities of the hybrid RGB/NIR acquisition workflow can provide robust liveness detection without requiring additional hardware or significant change in the acquisition workflow. A detailed system description and suitable workflow are presented.

Initial proof-of-concept experiments have shown that the proposed technique is effective for detecting a range of presentation attacks including a sophisticated attack based on

a lifelike mannequin that duplicates the characteristics of a human face and eyes. The conclusions are that iris biometrics can be made more robust than other well-known biometrics such as fingerprint and face by taking advantage of the handheld acquisition flow and the latest visible/NIR CMOS image sensor technologies.

Future work will involve developing an improved database to verify this technique over an enlarged user group and introducing additional advanced attack scenarios.

REFERENCES

[1] S. Curtis, "Smartphone at 20: IBM Simon to iPhone 6," *The Telegraph*, Aug. 2014.

[2] S. Curtis, "Quarter of the world will be using smartphones in 2016," *The Telegraph*, Dec. 2014.

[3] H. Orman, "Did you want privacy with that?: personal data protection in mobile devices," *IEEE Internet Comput.*, vol. 17, no. 3, pp. 83–86, May. 2013.

[4] D. Siewiorek, "Generation smartphone," *IEEE Spectr.*, vol. 49, no. 9, pp. 54–58, Aug. 2012.

[5] A. Smith, "U.S. Smartphone Use in 2015," Pew Research Centre, Apr. 2015.

[6] N. L. Clarke and S. M. Furnell, "Authentication of users on mobile telephones - A survey of attitudes and practices," *Comput. Secur.*, vol. 24, no. 7, pp. 519–527, Oct. 2005.

[7] P. Corcoran, "Biometrics and consumer electronics: a brave new world or the road to dystopia?" *IEEE Consum. Electron. Mag.*, vol. 2, no. 2, pp. 22–33, Apr. 2013.

[8] C. Sousedik and C. Busch, "Presentation attack detection methods for fingerprint recognition systems: a survey," *IET Biometrics*, vol. 3, no. 4, pp. 219–233, Dec. 2014.

[9] D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcão, and A. Rocha, "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 4, pp. 864–879, Apr. 2015.

[10] T. Ring, "Spoofing: are the hackers beating biometrics?," *Biometric Technol. Today*, vol. 2015, no. 7, pp. 5–9, Aug. 2015.

[11] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," *ACM Comput. Surv.*, vol. 47, no. 2, pp. 1–36, Jan. 2015.

[12] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Syst. Video Technol*., vol. 14, no. 1, pp. 4-20, Jan. 2004.

[13] J. Daugman, "How iris recognition works," *IEEE Trans. Circuits Syst. Video Technol*., vol. 14, no. 1, pp. 21–30, Jan. 2004.

[14] A. Martin, "NTT Docomo takes another step into a future without passwords," *Wall Street Journal, Japan,* May. 2015.

[15] S. Thavalengal, P. Bigioi, and P. Corcoran, "Iris authentication in handheld devices – considerations for constraint-free acquisition," *IEEE Trans. Consumer Electron.,* vol. 61, no. 2, pp. 245-253, May 2015.

[16] S. Thavalengal, I. Andorko, A. Drimbarean, P. Bigioi, and P. Corcoran, "Proof-of-concept and evaluation of a dual function visible/NIR camera for iris authentication in smartphones," *IEEE Trans. Consumer Electron.,* vol. 61, no. 2, pp. 137-143, May 2015.

[17] A. Czajka, "Pupil dynamics for iris liveness detection," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 4, pp. 726–735, Feb. 2015.

[18] B. Toth, "Liveness detection: iris," *Encyclopedia of Biometrics*, Springer US, pp. 931–938, 2009.

[19] Z. Akhtar, C. Micheloni, and G. L. Foresti, "Biometric liveness detection: challenges and research opportunities," *IEEE Secur. Priv.*, vol. 13, no. 5, pp. 63–72, Sep. 2015.

[20] K. B. Raja, R. Raghavendra, and C. Busch, "Video presentation attack detection in visible spectrum iris recognition using magnified phase information," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 10, pp. 2048–2056, Oct. 2015.

[21] S. Thavalengal, P. Bigioi, and P. Corcoran, "Evaluation of combined visible/NIR camera for iris authentication on smartphones," *in Proc. IEEE International Conference on Computer Vision and Pattern Recognition Workshops*, Boston, USA, pp. 42–49, Jun. 2015.

[22] S. Thavalengal, P. Bigioi, and P. Corcoran, "Efficient segmentation for multi-frame iris acquisition on smartphones," *in Proc. IEEE International Conference on Consumer Electronics*, Las Vegas, USA, pp. 202–203, Jan. 2016.

[23] I. Bacivarov, M. Ionita, and P. Corcoran, "Statistical models of appearance for eye tracking and eye-blink detection and measurement," *IEEE Trans. Consumer Electron.,* vol. 54, pp. 1312–1328, Aug. 2008.

[24] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.

[25] E. Steinberg, P. Corcoran, A. V Smith, B. M. Mehta, and M. Seth, "Digital image processing composition using face detection information," *US Patent US 8989453 B2*, Mar.2015.

[26] D. M. Endres and J. E. Schindelin, "A new metric for probability distributions," *IEEE Trans. Inf. Theory*, vol. 49, no. 7. pp. 1858–1860, Jun. 2003.

[27] Working Group, "ISO/IEC 30107 - Information technology - biometrics - presentation attack detection," Int. *Stand. Ed.*, 2013.

[28] T. Shejin and A. Sao, "Significance of dictionary for sparse coding based face recognition," *in Proc. International Conference of the Biometrics Special Interest Group*, Darmstadt, Germany, pp. 1–6, Sep. 2012.

[29] S. Mandal, S. Thavalengal, and A. K. Sao, "Explicit and implicit employment of edge-related information in super-resolving distant faces for recognition," *Pattern Anal. Appl.*, pp. 1–18, Aug. 2015.

[30] T. N. Cornsweet and H. D. Crane, "Accurate two-dimensional eye tracker using first and fourth Purkinje images," *J. Opt. Soc. Am.*, vol. 63, no. 8, pp. 921–928, Aug. 1973.

[31] E. C. Lee, K. R. Park, and J. Kim, "Fake iris detection by using purkinje image," Advances *in Biometrics*, vol. 3832 no. 1, pp. 397–403, Jan. 2006.

[32] P. M. Corcoran, P. Bigioi, and F. Nanu, "Detection and repair of flash-eye in handheld devices," *in Proc. IEEE International Conference on Consumer Electronics*, Las Vegas, USA, pp. 213–216, Jan. 2014.

[33] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag New York, 2006.

BIOGRAPHIES

Shejin Thavalengal (S'13) is a graduate student member of IEEE and works with FotoNation Ltd and National University of Ireland, Galway. He received MS (by Research) degree from Indian Institute of Technology Mandi, India in 2013 and B.Tech from Kannur University in 2010. His research interests include biometrics and pattern recognition.

Tudor Nedelcu (S'15) is a graduate student member of IEEE and works with FotoNation Ltd and National University of Ireland, Galway. Tudor received his B.S and M.Sc. degree from University "Politehnica" of Bucharest in 2011 and 2013 respectively. His research interests include multispectral imaging and biometrics.

Petronel Bigioi (M'99-SM'05) is a senior member of IEEE, SVP of Engineering in FotoNation Ltd and a lecturer in College of Engineering & Informatics at NUI Galway. He is co-inventor on 200 granted US patents. His research interests include VLSI design, digital imaging, communication network protocols and embedded systems.

Peter Corcoran (M '95-F'10) is a Fellow of IEEE and a Professor at NUI Galway. His research interests include biometrics and consumer electronics and he is a board member of the IEEE Biometrics Council. He is co-author on 250+ technical publications and co-inventor on more than 250 granted US patents. In addition to his academic career, he is also an occasional entrepreneur, industry consultant and compulsive inventor.