Video Presentation Attack Detection in Visible Spectrum Iris Recognition Using Magnified Phase Information

Kiran B. Raja*†, R. Raghavendra*, Christoph Busch*†
{kiran.raja; raghavendra.ramachandra; chrishtoph.busch} @hig.no
*Norwegian Biometrics Laboratory, Gjøvik University College, 2802 Gjøvik, Norway
†Hochschule Darmstadt - CASED, Haardtring 100, 64295 Darmstadt, Germany

***Abstract*—The gaining popularity of the visible spectrum iris recognition has sparked the interest in adopting it for various access control applications. Along with the popularity of visible spectrum iris recognition comes the threat of identity spoofing, presentation or direct attack. This work presents a novel scheme for detecting video presentation attacks in visible spectrum iris recognition system by magnifying the phase information in the eye region of the subject. The proposed scheme employs modified Eulerian Video Magnification (EVM) to enhance the subtle phase information in eye region and novel decision module to classify it as artefact(spoof attack) or normal presentation. The proposed decision module is based on estimating the change of phase information obtained from EVM, specially tailored to detect presentation attacks on video based iris recognition systems in visible spectrum. The proposed scheme is extensively evaluated on the newly constructed database consisting of 62 unique iris video acquired using two smartphones - iPhone 5S and Nokia Lumia 1020. We also construct the artefact database with 62 iris acquired by replaying normal presentation iris video on iPad with retina display. Extensive evaluation of proposed presentation attack detection (PAD) scheme on the newly constructed database has shown an outstanding performance of $Average\ Classification\ Error\ Rate(ACER) = 0\%$ supporting the robustness of the proposed PAD scheme.**

***Index Terms*—Replay attack, Iris recognition, Visible Iris, Presentation attack detection, Spoof, Biometrics**

I. Introduction

With the improvement of the cameras in smartphones to produce high quality images, one can easily adapt it to perform iris recognition in visible spectrum. Along with the improved capabilities of the smartphones as biometric sensor, it has to be noted that the systems can be easily spoofed by presenting the source video. The source video captured in visible spectrum can be accessed from various sources such as social media. The threat to systems employing biometric characteristics increases in the event of source data being available to the person trying to attack the system [1]. Once the data is available, it can be used by any imposter to attack the authentication system to gain the access into secured environment. The presentation attacks on biometric systems may employ photo attack or video attack. Many biometric systems which employ video based authentication have been robustly built to detect the liveness of the subject by analyzing the motion [2], [3], [4], [5].

However, the problem becomes more complicated when the imposter employs video which is similar to source video to attack the system. The sophisticated video presentation attack is carried out by playing the source video in front of a biometric system is challenging due to the fact the replayed video consists of motion present in original video. Recent works have investigated such vulnerabilities for face based biometric systems by replaying the source or original video [6]. In order to differentiate the motion under normal presentation of subject and the motion observed due to video presentation attack, earlier works have employed EVM [6]. The magnified videos were further processed using series of operations that include extraction of histogram of optical flow vectors followed by Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).

At the same time, even though iris biometrics is more reliable, one cannot under estimate the possibility of attack on iris recognition systems. The problem has been well addressed in Near-Infra-Red iris imaging for the print attacks [7], [8]. The print attacks are successfully detected by quantifying the quality artefacts of image and ocular features of the presented image [7], [8], [9]. However, there has been no works reported earlier on spoof detection in visible spectrum iris recognition to the best of our knowledge. An unexplored avenue in the visible spectrum iris recognition is video presentation attacks, especially on smartphone platform. With this motivation, in this work we contribute in three folds and they can be summarized as:

- This is the first comprehensive work investigating video based iris recognition in visible spectrum on smartphone platform by employing two latest smarphones - iPhone 5S and Nokia Lumia 1020.
- Proposes a robust scheme for video presentation attack detection (PAD) by employing phase information obtained from video of eye region. The proposed scheme is validated with extensive experiments to test the robustness and applicability.
- This work contributes to non-profit research work by distributing our newly constructed iris video database - VSSIRISV free of cost. The newly constructed iris video database consists of 62 unique iris acquired using two new smartphones - iPhone 5S and Nokia Lumia 1020.

The collected iris database is further used to generate high quality video presentation attack. This is a major contribution of this work as there are no publicly available iris video databases acquired in the visible spectrum using smartphones.

The rest of the paper is organized as follows: Section II provides the details of the newly constructed iris video database. Section III presents the proposed technique to detect video presentation attack in visible spectrum iris recognition system. Section IV details the experimental protocols and the obtained results in this work. Section V provides the conclusive remarks and the possible future work in this direction.

II. Database

With the advanced cameras and functionality in smartphones, there is a new interest in using them for iris biometrics in visible spectrum [10]. The possibility of using iris image/video from smartphone platform for authentication in specific applications such as owner authentication for the device, banking applications is getting more realistic in the recent days. Although there is a visible spectrum iris image database available from BIPLab [11], there is no other iris video database captured using smartphone in visible spectrum. In order to identify the challenges of detecting the presentation attack in visible spectrum iris recognition systems, in this work we have constructed a new iris video database using two new smartphones - iPhone 5S and Nokia Lumia 1020. The newly constructed iris video database, Visible Spectrum Smartphone Iris Video (VSSIRISV) database consists of 62 unique iris captured from 31 different subjects. The iris video were captured by placing the smartphones at a distance of 15 to 20 inches from the subjects. The capture process was generalized to replicate the real-life scenario of unconstrained conditions with mixed illumination consisting of natural light and artificial room light. The duration of videos in the database vary from 2-4 seconds. The subjects were allowed to blink in natural intervals making it unconstrained iris acquisition. VSSIRISV database consists of 12 brown iris, 12 gray iris and 38 blue/green iris acquired from 11 female subjects and 20 male subjects. The statistics of the database is provided in the Figure 1.

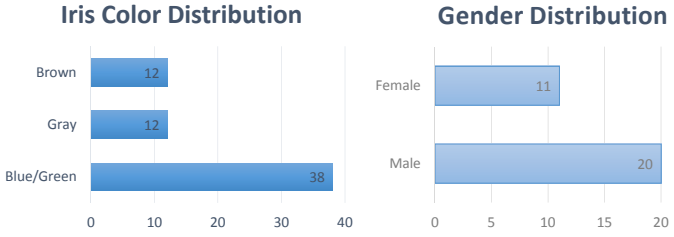

Fig. 1: VSSIRISV database statistics

A. Preprocessing of Live Iris Video Database

Due to large field of view, the acquired image from smartphone camera consists background along with the eye region and partial face. As this work intends to use the information provided by the eye region only, we eliminate all the other information present in the image by cropping the eye region. Based on the robust performance of Haar cascade based object detector [12], we use it to detect the eye region in the frames. The bounding box of detected eye region in the first frame is propagated to all the other frames in the video to obtain eye region from all the frames in a video.

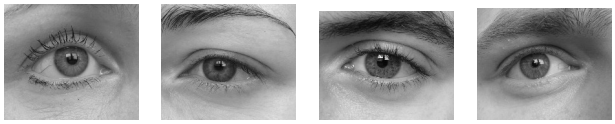

Fig. 2: Iris video frames captured using iPhone 5S

Fig. 3: Iris video frames captured using Nokia 1020

Figure 2 and Figure 3 present the gray scale version of detected eye region from the source video. As a biological organ, human eye is prone to blink at regular intervals. These intervals or frequency of blinking varies individually from person to person. As we do not want to use blink information to detect liveness, but the phase variation information from eye region, we manually inspect the iris video to obtain the frames between the blink. Thus, in this work, we have preprocessed the VSSIRISV database manually to obtain iris video consisting of 30 frames between the blinks which roughly corresponds to one second. These 30 frames are further processed to determine the liveness of the subject.

Another important feature of the captured database is that it is captured in unconstrained manner to simulate real-life scenario which has resulted in the involuntary head motion. As we do not intend to use the involuntary head motion in our work, it has to compensated/annulled to obtain unbiased results which are not due to motion of the head. Thus, we register all the preprocessed video without blinks using the phase correlation [13]. To register the frames based on phase correlation, we consider the first frame as reference and compute phase correlation for subsequent frames. The frames are aligned based on the phase correlation and final video consisting of overlapping regions is generated. Henceforth, the live iris video refers to the preprocessed video with registration and no-blink information. These videos are part of live video iris database and are used in the experimental evaluation.

B. Spoof Database Construction

This section provides the details of the construction of spoof or video presentation attack database. This being the first work on smartphone based visible spectrum video iris recognition, there is no public database for video attack samples. Thus, in this work we introduce a new high quality spoof iris database consisting of iris videos that can be used for presentation attacks. To generate the high quality spoof samples, we replay the live iris videos on the iPad with high quality display (retina-display). The replay videos were

TABLE I: decomposition of VSSIRISV database into Development and Testing set

Phone	Live Video				Artefact (Spoof) Video			
	Development		Testing		Development		Testing	
	Unique Eye	Samples per Eye	Unique Eye	Samples per Eye	Unique Eye	Samples per Eye	Unique Eye	Samples per Eye
iPhone 5S	10	2	52	2	10	2	52	2
Nokia 1020	10	2	52	2	10	2	52	2

carefully captured at a distance of the 9 inches from the display device in controlled manner with zero-influence of the external illumination to avoid the effects of the reflection and glare which may detriment quality of attack videos.

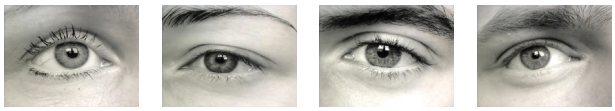

Fig. 4: Presentation attack video frames captured using iPhone 5S

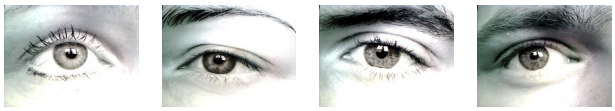

Fig. 5: Presentation attack iris video frames captured using Nokia 1020

The obtained videos were further preprocessed to remove edges contributed by the replay device. Thus, the final spoof videos consist of only the eye region obtained from the replay video with no additional context information such as background during capture, edges of the replay device and so on.

The correspondence of high quality spoof and live iris videos in this database provides a challenging database and thus aids us in obtaining unbiased comparison of the proposed technique. Figure 4 and Figure 5 present the sample frames from the spoof attack videos acquired by playing the live iris videos on iPad using iPhone 5S and Nokia Lumia 1020 smartphones. It can be observed that the video frames in live and spoof videos shown in the Fig. 2, Fig. 3 and Fig. 4, Fig. 5 are highly identical in terms of the quality.

C. Database Protocol

In order to effectively evaluate the proposed scheme, we divide the database in two sets, namely - development (10 unique iris) and testing set (52 unique iris). As the proposed technique does not use any training approaches, we have not considered the division of database to have training set. The development database consists of 10 unique live iris captured using iPhone 5S and Nokia Lumia 1020. Each unique iris in the development database is captured in 2 different instances. Thus the development database consists of 20 instance of iris videos from iPhone 5S and 20 instance of corresponding videos from Nokia 1020. The testing database consists of 52 unique iris videos captured from both the smartphones. Each unique iris in testing database has 2 video samples, thus 104 iris samples for each smartphone. Out of two videos for each eye, one video corresponds to reference and other to probe. Similarly, the presentation attack iris video corresponding to the live iris video are divided into development (10 iris video from each phone with 2 samples for each unique iris) and testing (52 iris video from each phone with 2 samples for each unique iris) set. The total decomposition of database in terms of testing and development set for the experiments is provided in the Table I. All the videos are first analyzed for the liveness using the proposed technique. Once the video is classified as a normal presentation, we use the frames from that video to obtain the verification score following the standard biometric system.

D. Availability of Database

In the view of limited availability of database for research on presentation attack detection in video based visible spectrum smartphone iris recognition, we intend to make the database available for non-profit research and academic purposes. Database can be obtained from www.nislab.no/biometrics_lab/vssirisv_db.

III. Proposed PAD Scheme

The proposed scheme for presentation attack detection is shown in the Figure 6. The iris video captured is preprocessed to consist of 30 frames between the blink interval. The frames are registered and aligned in the first step. The registered iris video is decomposed using Fourier Transform to obtain the phase and magnitude component. The phase component of the iris video is magnified to emphasize the variations in phase using modified EVM. The phase magnified video is analysed to make the decision. The techniques involved in deciding the iris video as presentation attack video is discussed in the upcoming sections.

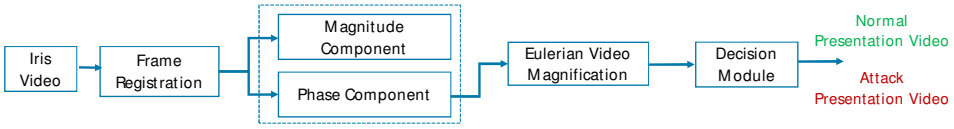

Fig. 6: PAD scheme for smartphone based visible spectrum iris recognition

A. Modified EVM for PAD

EVM can magnify the temporal variations in videos by decomposing it spatially and applying temporal filter [14]. In the case of presentation attack, the replayed video presents different information that can be attributed to the frequency of electronic screen. However, such information is very subtle and is hard to detect. If such information is magnified using linear magnification approaches in spatial domain, the resulting videos contain larger amount of noises. An alternative approach is to employ phase information to determine the presentation attacks. Phase based approaches provide robust performances which are not sensitive to noises as in the case of amplitude based approaches [15]. Another important intuition in using the phase information is to detect the additional frequency emitted when the video is played from an electronic screen. Inspired by the success of EVM in various cases to magnify the motion [14], [16], in this work, we have adopted modified version of EVM to enhance the small variations in phase component of video frame. The modified version of EVM uses the phase information as input. The video frame is decomposed using Fourier transform to obtain the phase information which is fed to EVM. The decomposed phase information is spatially filtered using Laplacian pyramids and temporally filtered using the Butterworth lowpass filter to magnify the variations in phase of each frame in the video. The enhanced phase variation in the video is used to estimate the liveness of subject as normal presentation or attack presentation. The proposed algorithm devised to detect the presentation attack is presented in detail in the section below.

B. Estimation of Liveness Score for PAD

Given the phase enhanced video consisting of N number of frames obtained using modified EVM, we perform a series of operations to detect the presentation attack. Since, performing computations on each of the frame is expensive in terms of memory and speed, each frame is downscaled to a smaller size. In this work, we have employed a downscaling size of 100×100 pixel based on the experimental trials on the development database. The magnified phase variation of each frame is normalized to have the values in the range of $0\ to\ 1$. Let F be the magnified phase variation of frame obtained from the EVM, then j^{th} normalized frame $NorF(j)$ is given by:

$$NorF(j) = \frac{F(j)}{max(F(j))} \quad \text{where } j = 1 : N \tag{1}$$

The normalized frame is further divided into non-overlapping blocks of specific size, $bx \times by$ as shown in Figure 7. We have employed a block size of 20×20 in our work based on the experimental trials on development database. This results in k number of blocks and thus each frame results in $k = 25\ blocks$ in our work. The normalized phase information of the block is further referred as normalized block phase variation and is represented as $NorFB(j)_k$ corresponding to j^{th} frame.

To effectively identify the presentation attack, we employ the sliding window approach with 6 frames, out of which 5 previous frames are used for making a decision on the present frame. The sliding window is propagated until 30 frames by incrementing one frame at a time. The sliding window is used to detect the rate of the change in the phase with respect to time. The size of the window was chosen based on the experimental trials conducted on development database. As illustrated in the Figure 7, for any given present j^{th} frame $NorF(j)$, differential phase variation for a block k is computed using 5 previous frames $NorF(j-1)\ to\ NorF(j-5)$. For a particular block k, the differential phase information with respect to 5 previous frames is given by:

$$\begin{aligned} DPI(j-5)_k &= NorFB(j)_k - NorFB(j-5)_k \\ DPI(j-4)_k &= NorFB(j)_k - NorFB(j-4)_k \\ DPI(j-3)_k &= NorFB(j)_k - NorFB(j-3)_k \\ DPI(j-2)_k &= NorFB(j)_k - NorFB(j-2)_k \\ DPI(j-1)_k &= NorFB(j)_k - NorFB(j-1)_k \\ &\text{for } k = 1, 2, \cdots 25 \end{aligned} \tag{2}$$

The final differential phase variation for a particular block in a frame j is obtained by determining the maximum of all the computed differences given by Equation 2.

$$\begin{aligned} DPI(j)_k &= max\{DPI(j-5)_k, \cdots DPI(j-1)_k\} \\ &\text{for } k = 1, 2, \cdots 25 \end{aligned} \tag{3}$$

The cumulative phase information, CPI is obtained for the entire frame j by summing all differential phase information across all the blocks b in the frame j.

$$CPI(j) = \sum_{x=1}^{k} DPI(j)_x \tag{4}$$

The cumulative phase information given by Equation 4 is computed in a similar manner for all the frames by employing the sliding window with 6 frames as described earlier.

In order to have the obtained score mapped to fixed interval values, the cumulative phase information is further normalized to a value between $0.5\ to\ 1$. The normalized cumulative phase information is used to determine the presentation attack. If the obtained value is above the threshold, the video is classified as a presentation attack. For a set of obtained CPI corresponding to a particular frame j, we apply single sided logistic or

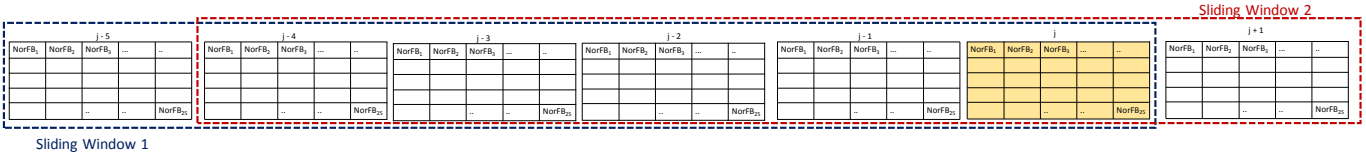

Fig. 7: Schematic of the block based phase variation information

sigmoid function to obtain normalized CPI represented as $NCPI$:

$$NCPI(j) = \frac{1}{1 + \exp -CPI(j)} \tag{5}$$

Based on the experimental trials on development database, a threshold value of 0.7 is obtained. The normalized cumulative phase information is thresholded against a value indicated by $Th = 0.7$ to obtain the liveness score LS for a particular frame j.

$$LS(j) = \begin{cases} 1, & \text{if } NCPI(j) \leq Th \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

The frames with $LS = 1$ are classified as normal presentation or live subjects and frames with other values are classified as the presentation attacks videos. The obtained liveness scores at various frames can be used to decide on the presentation category as normal presentation or attack presentation.

IV. Experiments and Results

The proposed scheme to detect the presentation attack is extensively evaluated on our newly constructed smartphone based video iris database acquired in visible spectrum. As discussed in the Section II, the VSSIRISV database consists of 62 unique live iris video and 62 unique attack iris videos from two different smartphones. Under the assumption that presentation attack videos can be obtained from various sources, we generate the attack videos from live iris videos obtained from iPhone 5S and Nokia 1020. Both of these videos are replayed on the high quality display device to attack the biometric system based on either iPhone 5S or Nokia 1020. This gives rise to a scenario where live video is obtained from a particular sensor, say iPhone 5S, and the video from the same smartphone can be used to attack the system by presenting it. An alternative case is where the live iris video and spoof iris video correspond to videos originating from different smartphones. The second case is very important due to the fact people tend to change the smartphones quite often due to availability of better features, reduced cost and limited shelf life. In order to assess the robustness of proposed system, we consider both situations and propose two different protocols to evaluate the proposed PAD scheme.

A. Protocol 1

In the protocol 1, the biometric system is attacked by the videos originating from same smartphone. In accordance to this protocol, the biometric system employing iPhone 5S is challenged by the attack videos by replaying the videos from iPhone 5S. In terms of similar arguments, the system employing Nokia 1020 are attacked using the replay videos of Nokia 1020. This protocol takes care of the attacks based on the same sensors. Since the live video and presentation attack video originating from same sensor are highly identical in terms of quality, this protocol intends to gauge the robustness of the proposed technique.

B. Protocol 2

Under the assumption that the imposter can use iris video obtained using different smartphone to attack the smartphone based visible spectrum iris biometric system, we propose to evaluate a situation where the attack videos and reference videos originate from different smartphones. Thus, in a system employing iPhone 5S as the biometric sensor, the attack video corresponding to Nokia 1020 is replayed and vice-versa. This protocol evaluates the reliability of the proposed technique to address the cross sensor presentation attacks in smartphone based visible spectrum iris recognition.

C. Experimental Set-up

Our newly constructed VSSIRISV database is employed in this work to evaluate the proposed scheme. The videos are used to determine the liveness of the subject as normal presentation or attack presentation using the proposed scheme mentioned in Section III. The various parameters such as the threshold for determining the liveness is based on the development database consisting of 10 unique iris videos with 2 instances for each unique iris captured from 2 different smartphones. The testing database from VSSIRISV database consisting of 52 unique iris videos from iPhone 5S and Nokia 1020 each are employed in the evaluation of the work. Each of the unique iris has 2 video instances which correspond to reference and probe video. Similarly, 52 unique iris spoof videos from iPhone 5S and Nokia 1020 are employed in the experimental evaluation.

All the results related to the proposed scheme of presentation attack detection have been disclosed in terms of Attack Presentation Classification Error Rate (APCER) and Normal Presentation Classification Error Rate (NPCER) [17]. APCER is defined as the proportion of attack presentations incorrectly classified as normal presentations in a specific scenario while NPCER is defined as the proportion of normal presentations incorrectly classified as attack presentations in a specific scenario [17]. Further, we also disclose the results in terms of Average-Classification-Error-Rate (ACER) which is described as the average of APCER and NPCER. ACER is defined by the Equation 7 as:

$$ACER = \frac{APCER + NPCER}{2} \tag{7}$$

With a fixed threshold of $Th = 0.7$ to obtain the liveness score, in this work we have achieved ACER of 0%. The obtained threshold on development database is well suited for the decision after frame number 11. The same threshold when applied on the testing database, we obtain the $ACER = 0\%$ indicating general applicability of proposed scheme for presentation attack detection on video based smartphone iris recognition in visible spectrum.

D. Results

The Table II presents various ACER obtained on testing database when different frames starting from 6 till 11 are considered with a threshold of $Th = 0.7$. The results are indicated from frame number 6 as the frames $1\ to\ 5$ are used to

TABLE II: Presentation classification error rates with a Normalized Cumulative Phase Information (NCPI) threshold of 0.7. Note: Frame number 1 to 5 are used to make the decision on frame number 6 in our proposed approach.

Reference Video	Attack Video	Classification Error Rate (%)																	
		Frame 6			Frame 7			Frame 8			Frame 9			Frame 10			Frame 11		
		APCER	NPCER	ACER	APCER	NPCER	ACER	APCER	NPCER	ACER	APCER	NPCER	ACER	APCER	NPCER	ACER	APCER	NPCER	ACER
Nokia	Nokia	100.00	100.00	100.00	100.00	98.07	99.04	100.00	75.00	87.50	50.00	19.23	34.62	3.85	0.00	1.92	0.00	0.00	0.00
	iPhone	100.00	100.00	100.00	100.00	96.23	98.12	100.00	73.60	86.80	51.92	18.48	35.20	0.00	0.00	0.00	0.00	0.00	0.00
iPhone	iPhone	100.00	98.07	99.04	100.00	92.31	96.15	100.00	78.85	89.42	55.76	13.46	34.61	0.00	0.00	0.00	0.00	0.00	0.00
	Nokia	100.00	97.40	98.70	100.00	89.70	94.85	100.00	73.62	86.81	51.92	17.30	34.61	1.92	0.00	0.96	0.00	0.00	0.00

make the first decision at frame 6. From the obtained results, the best possible and reliable frame for making a decision is frame number 11 which provides ACER of 0% for all cases. The determined liveness score remains constant after 11^{th} frame from all our experiments. It can be observed from Figure 8 that the liveness score obtained using the proposed scheme is robust after frame number 11 with $Th = 0.7$. For the sake of simplicity, we have illustrated one case where the reference iris videos are captured from iPhone 5S and presentation attack videos are obtained by replaying videos from iPhone 5S on iPad. Similar patterns can be seen on all of the other experimental trials with different combinations of live and presentation attack videos.

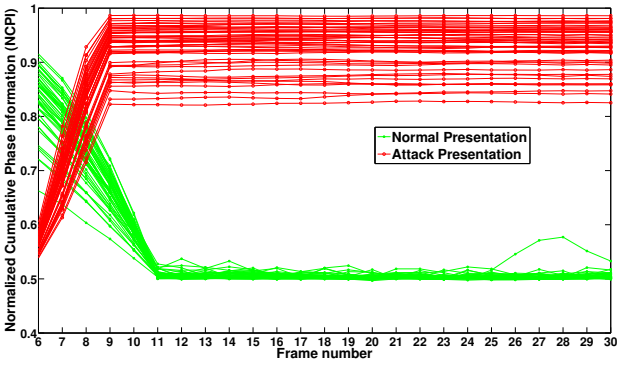

Fig. 8: Liveness score for iris video captured from iPhone 5S

E. Feature Extraction & Verification

Based on the decision made in terms of liveness score, a particular subject is further verified if the presentation is classified as normal presentation. Once a particular presentation is qualified as normal presentation, we perform the segmentation and normalization of the iris image for verification. Owing to the reported robustness of OSIRIS v4.1 [18], we have employed it for segmenting the iris images. Following the segmentation, we unwrap the iris pattern using Daugman's rubber sheet model [19]. Inspired by the success of the Local Binary Patterns and Sparse Representation Classifier (LBP-SRC) in the previously reported works [20], we have employed LBP-SRC in our work to obtain verification scores. The verification scores correspond to residual errors obtained using $L_1 - minimization\ via\ SPGL_1$ solver based on spectral gradient projection. In this work, we have adopted a protocol which uses 5 different frames from the reference video and 5 different frames from probe video to obtain the comparison score. The probe frame is iteratively swapped to obtain the comparison score and the all the 5 different comparison score is averaged to report the performance of the system.

1) Results of Protocol 1: Figure 9 provides the verification performance of the system without the attacks, with the attacks and with the proposed counter measure using PAD scheme. Figure 9 (a) indicates the performance of biometric system when 11^{th} frame is employed to make a decision for a system where reference videos are obtained from iPhone 5S and presentation attack videos are from iPhone 5S. Figure 9(b) indicates the performance of biometric system with reference videos captured from Nokia 1020 and the presentation attack videos are captured by videos from Nokia 1020. It can be observed that in both the cases the original performance of the biometric system is restored when frame number 11 is used for decision ($ACER = 0\%$). Thus, the proposed method is bale to detect the presentation attack when the source videos and presentation attack videos correspond to same smartphone.

2) Results of Protocol 2: This protocol is related to reference and presentation attack videos obtained from two different smartphones. Figure 9 (c) and (d) indicate the performance of the proposed scheme when the source video and attack video are generated from different smartphones. It can be observed from the Figure 9 (c) and (d) that the proposed system is able to detect the presentation attack successfully when frame number 11 is used to make decision. This result proves the adaptability of proposed PAD scheme irrespective of the smartphone employed in the video based visible spectrum iris recognition.

Additionally, Figure 10 presents the verification scores obtained for three groups of samples (valid iris videos (corresponding to normal presentation), imposters and presentation attackers) for the testing database. Figure 10(a) corresponds to scores of smartphone based visible spectrum iris recognition system using iPhone 5S and Figure 10(b) corresponds to Nokia 1020. It can be seen that the scores of the presentation attacks are in the same bins as of genuine scores which re-indicate the challenging samples in the database. The proposed technique is able to prevent the presentation attacks even under such high quality attack data.

V. Conclusion

With the gaining importance for visible spectrum iris recognition, in this work we have explored smartphone based video iris recognition in visible spectrum. The challenges of video based presentation attacks is explored in this work. Considering the high quality display devices capable of presenting attack videos with a nearly inseparable quality of difference, this work presents a novel scheme to detect the presentation

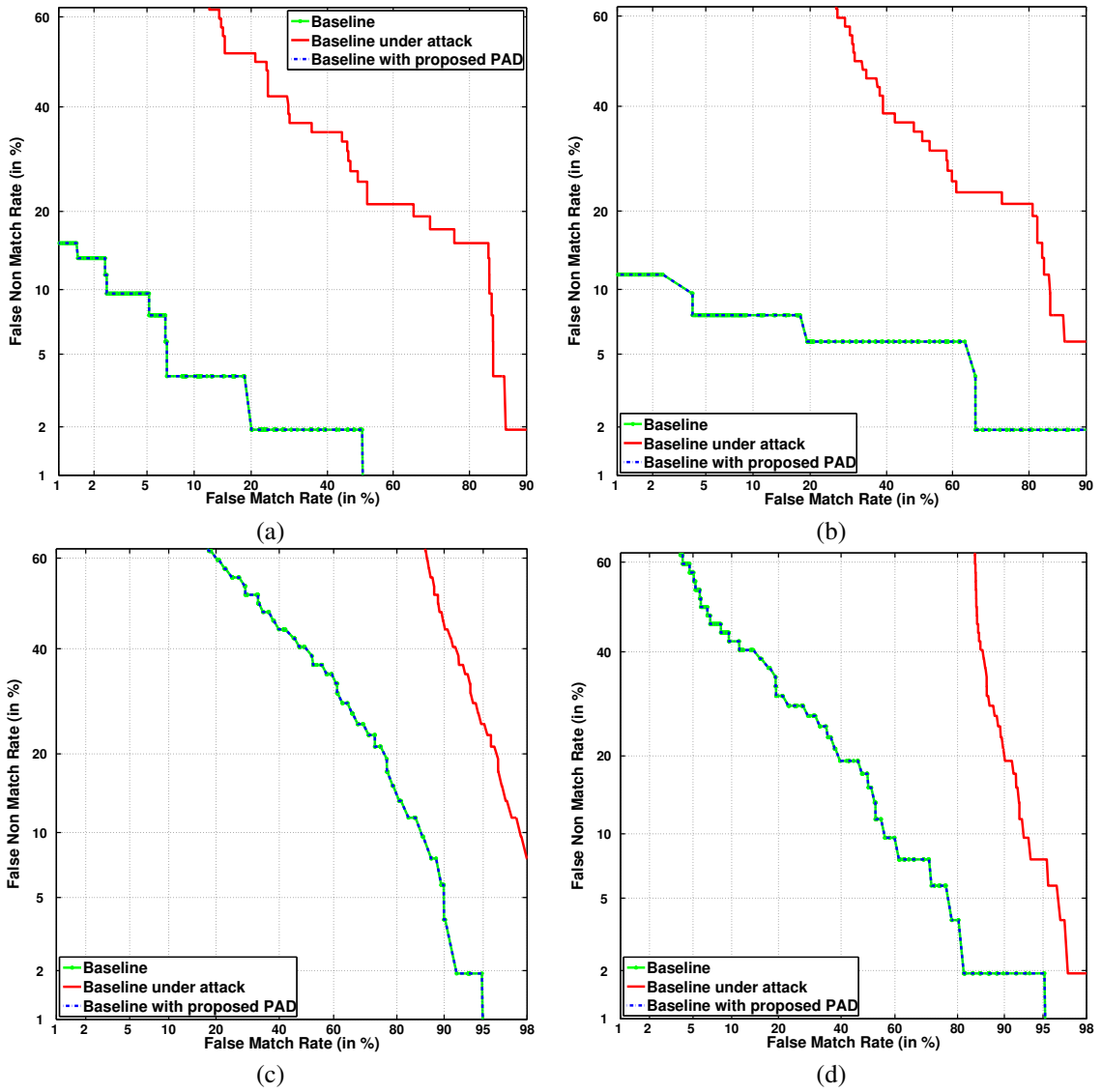

Fig. 9: (a) Baseline PAD for iPhone reference videos attacked with presentation attack videos generated from iPhone 5S videos after 11^{th} frame. (b) Baseline PAD for Nokia 1020 reference videos attacked with presentation attack videos generated from Nokia 1020 videos after 11^{th} frame. (c) Baseline PAD for iPhone 5S reference videos attacked with presentation attack videos generated from Nokia 1020 videos after 11^{th} frame. (d) Baseline PAD for Nokia reference videos attacked with presentation attack videos generated from iPhone 5S videos after 11^{th} frame.

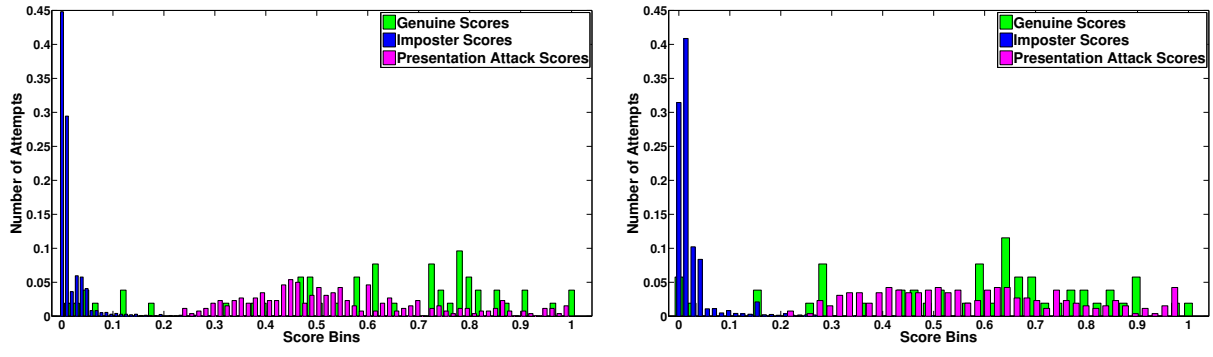

Fig. 10: (a) Score distributions for genuine, imposter and presentation attack iris videos for iPhone 5S (b) Score distributions for genuine, imposter and presentation attack iris videos for Nokia 1020

attack in visible spectrum iris recognition. The highlight of this work is in using smartphone as a biometric sensor for iris recognition and also addressing the high quality replay attack originating from electronic display. A database of 62 unique iris is captured using two new smartphones - iPhone 5S and Nokia Lumia 1020. The experiments are conducted to replicate a real-life scenario by collecting the data in unconstrained condition of mixed illumination under semi-cooperation from the subjects. The captured live iris videos are used to generate the presentation attack videos. In order to magnify the phase variations in the video of eye region, this work has employed modified EVM.

The obtained results from the experiments have indicated an ACER of 0%. The video presentation attacks can be identified as early as in 11^{th} frame. The proposed technique is experimentally validated for the reliable and robust performance using our newly constructed VSSIRISV database. It has been observed experimentally that the proposed counter-measure or PAD scheme brings the performance of system back the original performance level under normal presentations or no attacks.

Further, to support the research on smartphone based visible spectrum iris recognition using videos, we disclose the newly constructed VSSIRISV database for non-profit research and academic purposes. This database shall provide a basis for other researchers to propose new techniques for PAD in smartphone based visible spectrum iris recognition.

ACKNOWLEDGMENTS

The authors would like to extend their thanks to reviewers for their valuable comments that helped in improving the quality of the manuscript. The authors would also like to thank the numerous volunteers that contributed to make the data collection at our campus possible. Further, the authors wish to express thanks to Morpho (Safran Group) for supporting this work, and in particular to Morpho Research & Technology team for the fruitful technical and scientific exchanges related to this particular work.

REFERENCES

[1] A. B. Toth and J. Galbally, *Anti-Spoofing: Iris*. Springer-Verlag Berlin Heidelberg, 2005, pp. 970–977.

[2] A. Anjos, M. M. Chakka, and S. Marcel, "Motion-based counter-measures to photo attacks in face recognition," *Institution of Engineering and Technology - Biometrics*, Apr. 2013.

[3] A. Anjos and S. Marcel, "Counter-measures to photo attacks in face recognition: a public database and a baseline," in *2011 International Joint Conference on Biometrics (IJCB)*. IEEE, 2011, pp. 1–7.

[4] J. Maatta, A. Hadid, and M. Pietikainen, "Face spoofing detection from single images using texture and local shape analysis," *IET Biometrics*, vol. 1, no. 1, pp. 3–10, 2012.

[5] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, "A face anti-spoofing database with diverse attacks," in *2012 5th IAPR International Conference on Biometrics (ICB)*. IEEE, 2012, pp. 26–31.

[6] S. Bharadwaj, T. I. Dhamecha, M. Vatsa, and R. Singh, "Computationally efficient face spoofing detection with motion magnification," in *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 2013, pp. 105–110.

[7] J. Ortiz-Lopez, J. Galbally, J. Fierrez, and J. Ortega-Garcia, "Predicting iris vulnerability to direct attacks based on quality related features," in *2011 IEEE International Carnahan Conference on Security Technology (ICCST)*. IEEE, 2011, pp. 1–6.

[8] J. Galbally, J. Ortiz-Lopez, J. Fierrez, and J. Ortega-Garcia, "Iris liveness detection based on quality related features," in *2012 5th IAPR International Conference on Biometrics (ICB)*. IEEE, 2012, pp. 271–276.

[9] C.-W. Tan and A. Kumar, "Integrating ocular and iris descriptors for fake iris image detection," in *2nd International Workshop on Biometrics and Forensics, Malta*. IEEE, 2014.

[10] M. De Marsico, C. Galdi, M. Nappi, and D. Riccio, "Firme: Face iris recognition for mobile engagement," *Image and Vision Computing*, 2014.

[11] BIPLab, University of Salerno, "Mobile Iris CHallenge Evaluation (MICHE I and II)," http://biplab.unisa.it/MICHE/database/.

[12] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001.*, vol. 1. IEEE, 2001, pp. I–511.

[13] A. Wong and J. Orchard, "Robust multimodal registration using local phase-coherence representations," *Journal of Signal Processing Systems*, vol. 54, no. 1-3, pp. 89–100, 2009.

[14] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Transactions on Graphics (TOG)*, vol. 31, no. 4, p. 65, 2012.

[15] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.

[16] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, "Phase-based video motion processing," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, p. 80, 2013.

[17] ISO/IEC TC JTC1 SC37 Biometrics, *ISO/IEC CD 30107-1. Biometrics - Presentation Attack Detection - Part 1*, Framework. International Organization for Standardization and International Electrotechnical Committee, March, 2014.

[18] G. Sutra, B. Dorizzi, S. Garcia-Salicetti, and N. Othman, "A Biometric Reference System for Iris, OSIRIS version 4.1," 2012.

[19] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004.

[20] R. Raghavendra, Kiran B. Raja, B. Yang, and C. Busch, "Combining iris and periocular recognition using light field camera," in *2nd IAPR Asian Conference on Pattern Recognition (ACPR2013)*. IEEE, 2013.

[21] H. Zhang, Z. Sun, and T. Tan, "Contact lens detection based on weighted lbp," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 4279–4282.

[22] J. Galbally, S. Marcel, and J. Fierrez, "Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition," *Image Processing, IEEE Transactions on*, vol. 23, no. 2, pp. 710–724, 2014.

[23] R. Raghavendra and C. Busch, "Robust scheme for iris presentation attack detection using multiscale binarized statistical image features," *Information Forensics and Security, IEEE Transactions on*, vol. 10, no. 4, pp. 703–715, 2015.

Kiran Bylappa Raja received his Bachelor of Engineering degree in Electronics and Communication from Visvesvaraya Technological University, Belgaum, India, in 2007, and the M.S. degree from Erasmus Mundus Master CIMET, Saint-Etienne, France, and Gjvik, Norway, in 2013. From 2007 to 2011, he was with Tata Elxsi Ltd., Bangalore, India, Bally Technologies Ltd., Bangalore, India, and Technicolor Inc., CA, USA. He is currently pursuing the Ph.D. degree in computer science with Gjvik University College, Gjvik, and Hochschule Darmstadt, Darmstadt, Germany. His main research interests include statistical pattern recognition, image processing, and machine learning with applications to biometrics. He has authored several papers.

R. Raghavendra received the bachelors degree from the University of Mysore (UOM), Mysore, India, the masters degree in electronics and communication from Visvesvaraya Technological University, Belgaum, India, and the Ph.D. degree in computer science and technology from UOM and Institute Telecom, and Telecom Sudparis, Evry, France (carried out as a collaborative work). He is currently a Researcher with the Norwegian Biometric Laboratory, Gjvik University College, Gjvik, Norway. He was a Researcher with the Istituto Italiano di Tecnologia, Genoa, Italy. His main research interests include statistical pattern recognition, data fusion schemes and random optimization, with applications to biometrics, multimodal biometric fusion, human behavior analysis, and crowd behavior analysis. He has authored several papers, and is a reviewer for several international conferences and journals.

Christoph Busch received the Diploma degree from the Technical University of Darmstadt (TUD), Darmstadt, Germany, and the Ph.D. degree in computer graphics from TUD, in 1997. He joined the Fraunhofer Institute for Computer Graphics, Darmstadt, in 1997. He is a member of the Faculty of Computer Science and Media Technology with Gjvik University College, Gjvik, Norway, and holds a joint appointment with the Faculty of Computer Science, Hochschule Darmstadt, Darmstadt. Further, he lectures in biometric systems with DTU since 2007. His research includes pattern recognition, multimodal and mobile biometrics, and privacy enhancing technologies for biometric systems. He is the Co-Founder of the European Association for Biometrics and a Convener of WG3 in ISO/IEC JTC1 SC37 on Biometrics. He co-authored over 280 technical papers, and has been a Speaker at international conferences. He served for various program committees, and is also an appointed member of the Editorial Board of the IET Journal on Biometrics.