PRESENTATION ATTACK DETECTION ALGORITHM FOR FACE AND IRIS BIOMETRICS

R. Raghavendra *Christoph Busch*

Norwegian Biometric Laboratory, Gjøvik University College, Norway
Email: $\{raghavendra.ramachandra, christoph.busch\}$ $@hig.no$

ABSTRACT

Biometric systems are vulnerable to the diverse attacks that emerged as a challenge to assure the reliability in adopting these systems in real-life scenario. In this work, we propose a novel solution to detect a presentation attack based on exploring both statistical and Cepstral features. The proposed Presentation Attack Detection (PAD) algorithm will extract the statistical features that can capture the micro-texture variation using Binarized Statistical Image Features (BSIF) and Cepstral features that can reflect the micro changes in frequency using 2D Cepstrum analysis. We then fuse these features to form a single feature vector before making a decision on whether a capture attempt is a normal presentation or an artefact presentation using linear Support Vector Machine (SVM). Extensive experiments carried out on a publicly available face and iris spoof database show the efficacy of the proposed PAD algorithm with an Average Classification Error Rate (ACER) $= 10.21\%$ on face and $ACER = 0\%$ on the iris biometrics.

***Index Terms*—** Biometrics, Spoof, Attack detection, Face, Iris

1. INTRODUCTION

In recent years, biometric systems have been deployed in numerous security applications. As the adoption of biometric system increases their vulnerability to various presentation attack also gained momentum. The goal of the presentation attack is to subvert a biometric system by presenting a biometric artefact. Even though the vulnerability of the biometric system has been attested for all available modalities, the face and iris biometric system have proven to be more vulnerable [1–3]. This is because, no additional effort is required to generate either face or iris biometric artefacts as one can generate these artefacts by capturing a photo or video of the legitimate subject even without its notice or one can also obtain these images from the wide spread social media websites. Furthermore, generating face and iris artefacts is not only easy but also cost effective.

There exists numerous PAD techniques that can be broadly classified into three main groups namely: (1) Motion based schemes (2) Micro-texture based scheme (3) Image quality analysis schemes. The micro-texture and motion based schemes have proven their effectiveness especially for facial biometrics [4] [5]. The motion based schemes are more appropriate to video based presentation attacks where the idea is to analyze the abnormal motion either using motion correlation [6] or non-rigid motion analysis based on GMM [7] or Eulerian magnification [7] or dynamic texture analysis [7]. Most of the micro-texture schemes for face PAD belong to the class of Local Binary Patterns (LBP) [5] or to the filters based on Difference of Gaussians (DoG) [8] in addition to this, frequency analysis based on Fourier Transform is also addressed [7]. The image quality analysis for the face PAD algorithm involves analyzing the basic quality features like sharpness, contrast, and etc. While for the iris PAD algorithms basically involves analyzing the image quality. In [9], 25 different well established image quality measures are employed to detect the artefact (or fake) iris sample. While in [10] image quality measure like local and global contrast, frequency distribution rates in addition to Gray Level Cooccurrence matrix (GLCM) is employed that shows the best performance in the LivDet 2013 competition [3].

In this work, we present a novel PAD algorithm based on statistical features extracted using BSIF [11] and the Cepstral features extracted using 2D-Cepstrum [12]. The proposed PAD algorithm forms a generic solution for both face and iris biometric modality and thus different from the state-of-the-art PAD schemes. Given a biometric sample, the proposed PAD algorithm will extract both BSIF and 2D Cepstrum features separately, which are then fused to form a single feature vector before obtaining a decision using the linear SVM. Extensive experiments are carried out on the face and iris publicly available database. For face we employed CASIA face spoof database [8] by considering its variability not only in generating the artefacts but also its capturing protocols that includes three different cameras with varying resolutions. For the iris analysis we choose ATVS- Fake Iris database [13] that comprised of 50 subjects with 800 iris artefacts. We also present the comparative analysis of our proposed scheme with the well-established state-of-the-art PAD schemes in both face [8] [5] and iris [9] [13].

This work is funded by the EU 7th Framework Program (FP7/2007-2013) under grant agreement n^o 284862 for the large-scale integrated project FIDELITY.

The rest of the paper is organized as follows: Section 2 presents the proposed PAD algorithm, in Section 3 we discuss the results obtained on the proposed PAD scheme and conclusion is drawn in the Section 4.

2. PROPOSED PAD ALGORITHM

Figure1 shows the block diagram of the proposed PAD algorithm that employs two different feature extraction schemes namely 2D-Cepstrum [12] and BSIF [11]. We then fuse these two feature vectors by concatenation to form a single vector before obtaining a decision using linear SVM classifier as explained below.

2.1. Feature Extraction using 2D Cepstrum

The 2D Cepstrum based feature extraction is one of the most successful and widely used technique in the domain of speech and image processing. Thus, inspired by its wide success in various speech and image processing areas, we propose its exploration to assess the presentation attack detection in biometric systems. Most of the attacks are presented by creating a biometric artefact in the form of either a photo or replaying a video in front of capturing device. Hence these presented artefacts when captured by the devices (or cameras) will tend to exhibit a larger frequency components when compared to that of real biometric samples. Since these high frequencies can be more emphasized using 2D Cepstrum by exploring the non-uniform binning of the spectral information. It is our assertion that, the use of 2D Cepstrum features can present comprehensive information which in turn can be used to detect the presentation attacks on the biometric system.

Let $y(a, b)$ be the captured biometric sample, then the 2D Cepstrum can be computed by the following four steps. First, obtain the 2D Discrete Fourier Transform (DFT) of $y(a, b)$ as following:

$$Y(A, B) = \frac{1}{N} \sum_{a} \sum_{b} y(a, b) * e^{-j2\pi(ua+ub/N)} \tag{1}$$

In the next step, divide the obtained DFT data from Eq.1 into non-uniform bins in a logarithmic manner and compute the energy of each bin as follows:

$$|E(m, n)^2| = \sum_{k,l \in B(m,n)} w(k, l) Y(k, l) \tag{2}$$

Where, $B(m, n)$ is the $(m, n)^{th}$ cell of the logarithmic grid corresponding to the weight $w(k, l)$. Since bins are smaller at low frequencies when compared to higher frequencies, the use of 2D Cepstrum will emphasis the high frequencies. Finally, the 2D Cepstrum are computed using inverse 2D DFT as follows:

$$\hat{y}(p, q) = F_2^{-1}\left(log\left(|E(m, n)^2|\right)\right) \tag{3}$$

Where p, q denote the 2D Cepstral frequency coordinates and F_2^{-1} denotes the 2D inverse DFT.

Figure 2 and 3 illustrates the qualitative results of the 2D Cepstrum obtained on both real face and iris with its artefacts biometric samples. Here, it is interesting to observe that, the presence of artefact samples denote the increase in the strength of high frequency components (observed as the dark color in the Figure 2). This justifies the applicability of the proposed 2D Cepstum for accurate presentation attack detection.

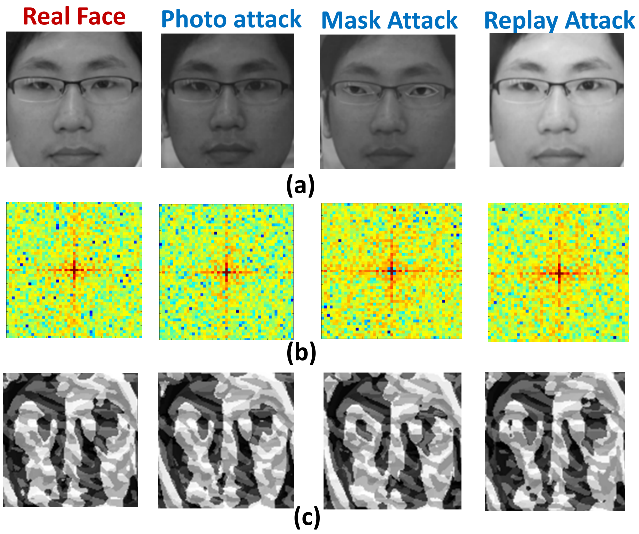

Fig. 2. Qualitative results of the proposed PAD algorithm on face (a) Raw images (b) 2D Cepstrum results (c) BSIF results

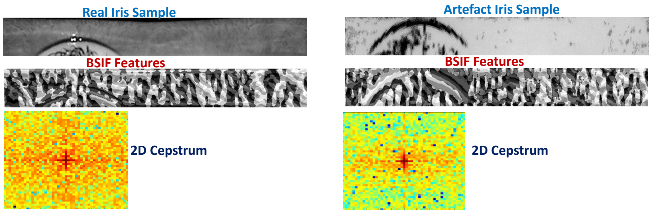

Fig. 3. Illustration of qualitative results obtained using proposed PAD algorithm on ATVS fake iris biometrics

2.2. Binarized Statistical Image Features (BSIF)

The idea of the BSIF is to represent each pixel as a binary code obtained by computing its response to a filter that are trained utilizing the statistical properties of the natural images. In this work, we employed the open-source filters [11] that are trained using 50000 image patterns randomly sampled from 13 different natural images [14]. The learning process to construct these statistically independent filters involves three

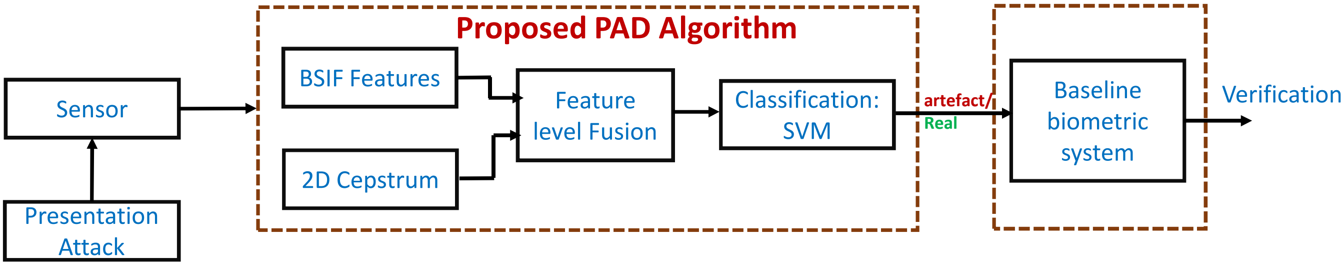

Fig. 1. General block diagram of the proposed spoof-resistant biometric system

main steps (1) Mean subtraction of each patch (2) Dimensionality reduction using Principle Component Analysis (PCA) (3) Estimation of statistically independent filters (or basis) using Independent Component Analysis (ICA). Thus, given a biometric sample $y(a, b)$ and a filter h_i , the filter response is obtained as follows [11]:

$$r_i = \sum_{x,y} y(a, b) h_i(x, y) \tag{4}$$

Where x and y denotes the size of the 2D face image patch and $h_i, \forall i = \{1, 2, \ldots, n\}$ denotes the number of linear filters whose response can be computed together which in turn is binarized to obtain the binary string as follows [11]:

$$b_i = \begin{cases} 1, & \text{if } r_i > 0 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Finally, the BSIF features are obtained as the histogram of pixel's binary codes that can effectively characterize the texture components in the 2D face/iris image, which are denoted as S_y.

Figure 2 - 3 show the qualitative results of the BSIF features and illustrate the presence of minor differences in texture and smoothness characteristics that were effectively captured using BSIF features. That further justifies our proposed BSIF features for this precise application.

2.3. Feature level fusion and classification

After obtaining both 2D Cepstral ($\hat{y}(p, q)$) and statistical (S_y) features, we perform the feature level fusion by carrying out feature concatenation to obtain a new feature vector $F_e = (\hat{y}(p, q)||S_y)$. We then employ the linear SVM classifier to determine whether the captured sample belongs to real or artefact. The SVM classifier is first trained using a set of positive (real faces) and negative (artefact) samples according to the standard protocol described for both face and iris database.

2.4. Baseline face and Iris system

In order to evaluate the performance of 2D face recognition system, we employ the well-established face recognition baseline system based on the Sparse Representation Classifier (SRC) [15]. While for the iris biometric, we employed the scheme based on the mean absolute deviation as proposed in Li Ma et al. [16].

3. EXPERIMENTAL RESULTS AND DISCUSSION

We evaluate the proposed PAD scheme on publicly available face and iris spoof databases. For face, we select the CASIA face spoof database [8] as it comprised of three different ways of generating face biometric artefacts namely photo, mask and replay attacks played using tablet (RA) from 50 subjects. In addition, the artefacts are recorded using three different resolution sensors that allows one to study the role of PAD algorithms across various variations due to camera resolution and artefacts when compared to other available databases. For iris biometric, we employed the ATVS fake iris database [13] that comprised of 50 subjects whose artefacts are generated by printing the real iris on the high quality paper before capturing with LG EOU3000 device.

Experimental results presented in this work are carried according to the protocol presented by the corresponding face and iris spoof databases. The performance of the proposed PAD algorithm is measured using two kind of errors [17] namely: (1) Attack Presentation Classification Error Rate (APCER) where attack (fake or artefact) presentation incorrectly classified as normal (real) presentation. (2) Normal Presentation Classification Error Rate (NPCER) where normal presentation incorrectly classified as attack samples. Finally, the performance of the overall PAD algorithm is presented in terms of Average Classification Error Rate (ACER) such that, $ACER = \frac{(APCER+NPCER)}{2}$ thus, the lower the ACER Values the better is the performance.

Table 1. Performance of the proposed PAD with varying resolution on CASIA face spoof database

Training Set	Performance on Testing Set in ACER (%)								
	Low Resolution (LR)			Middle Resolution (MR)			High Resolution (HR)		
	Photo	Mask	RA	Photo	Mask	RA	Photo	Mask	RA
LR	6.75	7.94	2.10	0.07	0.25	2.18	0.00	0.03	0.22
MR	3.95	3.58	2.75	4.40	7.05	4.62	0.07	0.03	0.10
HR	0.02	0.27	0.01	0.05	0.02	0.02	8.16	10.55	2.14

Table 1 shows the performance of the proposed PAD algorithm on the CASIA face spoof database. Here, we present a comprehensive analysis to understand the role of image resolutions (or interoperability to different cameras) especially on the presentation attack and its impact on the proposed PAD scheme. To this extent, we train the proposed PAD system with the samples captured in one resolution (or one camera) at a time i.e. for instance, we train the PAD system with Low resolution samples and measure its performance with the artefacts generated using all three kind of resolutions namely Low Resolution (LR), middle Resolution (MR) and High Resolution (HR). Based on the obtained results, it is interesting to observe that, the performance of the PAD algorithm shows the strong relationship with the artefacts generated with different resolutions. It can be observed that, when a PAD system is trained only with LR samples, it fails to detect the artefacts generated using the samples captured by the same camera (or same resolution) when compared to other camera resolutions. Thus the proposed PAD algorithm shows larger ACER values when trained with LR samples and tested with artefacts generated using LR samples when compared to the artefacts generated using MR and HR. In fact, we observed similar performance with the state-of-the-art PAD algorithms based on LBP [5] and also on DoG filters [8] for simplicity, we have not included in this work. This experiment strongly suggests that, the influence of attacking the biometric system does not depend upon how good the biometric artefact is generated rather depends on the camera characteristics especially the resolution. Thus, if the attacker can successfully generate the artefact from the face sample captured using similar kind of the camera used by the biometric system on which he is intended to attack, then the vulnerability of such a system is very high for these kind of attacks. Table 2 shows the quantitative performance of the proposed scheme on both face and iris spoof databases. In addition, we also present the comprehensive comparison of the proposed scheme with the well adopted state-of-the-art schemes in both face and iris attack detection. It can be observed from the Table 2 that, the proposed PAD algorithm outperformed the state-of-the-art schemes in both face and iris biometrics. The proposed PAD shows the best performance of $ACER = 10.21\%$ by improving the performance by 8% when compared with three different existing schemes on face biometrics. Further, the proposed PAD algorithm has demonstrated the impressive results of $ACER = 0\%$ on iris biometrics and emerged as robust and generic solution for presentation attack detection for this biometric modality.

Table 2. Comparative performance of the proposed PAD scheme on CASIA face and ATVS iris fake databases

Face		Iris	
Algorithms	ACER (%)	Algorithms	ACER (%)
LBP$^{u2}_{3\times3}$ - LDA [5]	21.01	IQA [9]	2.20
LBP$^{u2}_{3\times3}$ - SVM [5]	18.21	Quality features [13]	3.10
DoG - SVM [8]	26.72	GLCM [10]	5.60
Proposed Scheme	**10.21**	Proposed Scheme	**0.00**

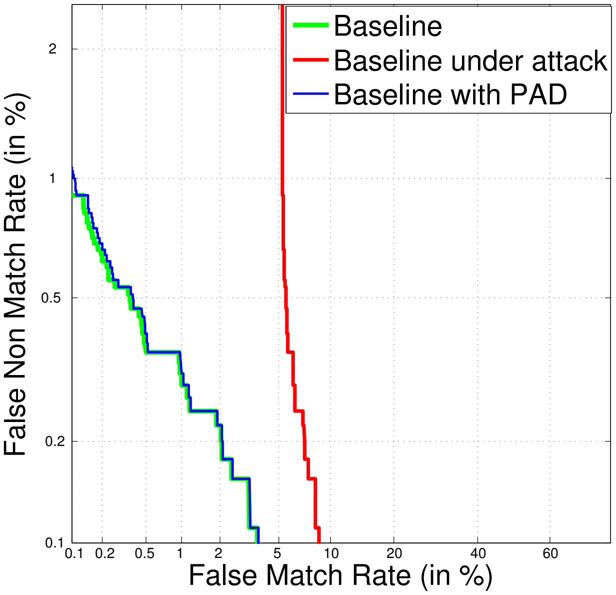

Fig. 4. Verification performance of the face biometrics system with and without proposed PAD algorithm

Figure 4 shows the verification performance of the face biometric system with and without our proposed PAD algorithm. It can be observed that, the presence of artefacts will drastically degrade the overall performance of the face biometric system by increasing both False Match Rate (FMR) and False Non Match Rate (FNMR) (as indicated by the red line in Figure 4). Further, it is also noted that, by adopting the proposed PAD algorithm we can improve the performance of the face biometric system to the nearly same performance where no attacks are presented (as shown in the green (baseline) and blue (baseline with proposed PAD)). Similar observation can also be noted for the iris biometric system as shown in the Figure 5. It is interesting to observe here that, the proposed PAD algorithm has completely discarded the attacks so that the overall system performance is undisturbed (as observed from the complete overlapping of green (baseline) and blue line (baseline with proposed PAD) in the Figure5. This strongly justifies the accuracy, robustness and applicability of the proposed PAD algorithm on both face and iris biometrics.

4. CONCLUSION

In this work, we presented a novel PAD algorithm that forms a generic solution for mitigating the attacks on both face and iris biometrics. The proposed method explores both micro-texture variation using Binarized Statistical Image Features (BSIF) and micro-frequency variations using 2D Cepstrum. We then combine these two features before obtaining the decision using linear SVM. Extensive experiments are carried

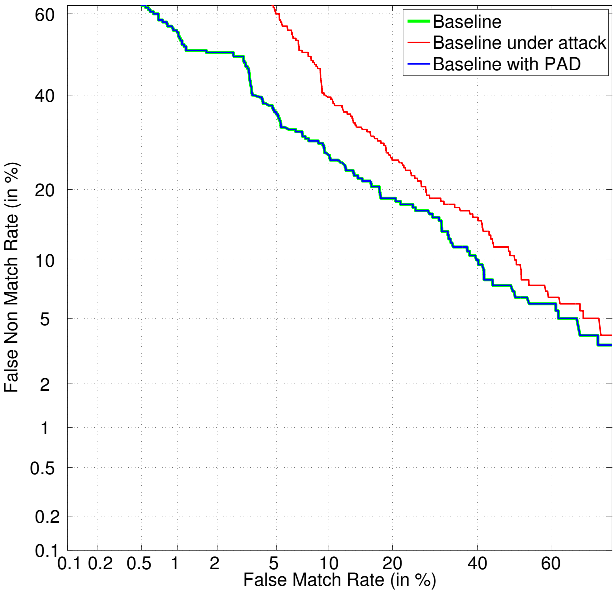

Fig. 5. Verification performance of the Iris biometric system with and without proposed PAD algorithm

out on the publicly available databases of face and iris biometrics. In addition, we present additional experiments that shows the performance of the proposed PAD algorithm with various camera resolution. Our experiments with various camera resolution especially on the face biometrics shows the sensitivity of the camera interoperability on the presentation attack detection. Further, experimental results also revealed that, the proposed PAD algorithm emerged as the best scheme with an $ACER = 10.21\%$ on face and $ACER = 0\%$ on the iris modality that further justifies the applicability of our proposed PAD algorithm in real-life scenarios.

REFERENCES

[1] K. A. Nixon, V. Aimale, and R. K. Rowe, "chapter: Spoof detection schemes," *Handbook of Biometrics*, vol. Springer-Verlag, 2008.

[2] Murali Mohan Chakka, Andre Anjos, Sebastien Marcel, Roberto Tronci, Daniele Muntoni, Gianluca Fadda, Maurizio Pili, Nicola Sirena, Gabriele Murgia, Marco Ristori, et al., "Competition on counter measures to 2-d facial spoofing attacks," in *International Joint Conference on Biometrics*, 2011, pp. 1–6.

[3] "Livdet-iris competition," `http://people.clarkson.edu/projects/biosal/iris/index.php`.

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

[5] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in *International Conference of the Biometrics Special Interest Group (BIOSIG)*, Sept 2012, pp. 1–7.

[6] André Anjos, Murali Mohan Chakka, and Sébastien Marcel, "Motion-based counter-measures to photo attacks in face recognition," *IET Biometrics*, 2013.

[7] Muhammad-Adeel Waris et al., "The 2nd competition on counter measures to 2d face spoofing attacks," in *International Conference of Biometrics (ICB)*, 2013, pp. 1–6.

[8] Zhiwei Zhang, Junjie Yan, Sifei Liu, Zhen Lei, Dong Yi, and S.Z. Li, "A face antispoofing database with diverse attacks," in *5th IAPR International Conference on Biometrics (ICB)*, March 2012, pp. 26–31.

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

[10] Ana F. Sequeira, Juliano Murari, and Jaime S. Cardoso, "Iris liveness detection methods in mobile applications," in *9th International Conference on Computer Vision Theory and Applications*, 2013, pp. 1–5.

[11] Juho Kannala and Esa Rahtu, "Bsif: Binarized statistical image features," in *Pattern Recognition (ICPR), 2012 21st International Conference on*, 2012, pp. 1363–1366.

[12] S. Cakir and A.E. Cetin, "Mel-cepstral methods for image feature extraction," in *17th IEEE International Conference on Image Processing (ICIP)*, Sept 2010, pp. 4577–4580.

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

[14] Aapo Hyvèarinen, Jarmo Hurri, and Patrick O Hoyer, *Natural Image Statistics*, vol. 39, Springer, 2009.

[15] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210 – 227, 2009.

[16] Lia Ma, Yunhong Wang, and Tieniu Tan, "Iris recognition based on multichannel gabor filtering," in *Proc. Fifth Asian Conf. Computer Vision*, 2002, vol. 1, pp. 279–283.

[17] ISO/IEC JTC1 SC37 Biometrics, *ISO/IEC WD 30107-3:2014 Information Technology - presentation attack detection - Part 3: testing and reporting and classification of attacks*, ISO, 2014.