Detection of Video-Based Face Spoofing Using LBP and Multiscale DCT

Ye Tian[1] and Shijun Xiang[1,2(✉)]

[1] School of Information Science and Technology, Jinan University, Guangzhou, China
[2] State Key Laboratory of Information Security,
Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
shijun_xiang@qq.com

Abstract. Despite the great deal of progress during the recent years, face spoofing detection is still a focus of attention. In this paper, an effective, simple and time-saving countermeasure against video-based face spoofing attacks based on LBP (Local Binary Patterns) and multiscale DCT (Discrete Cosine Transform) is proposed. Adopted as the low-level descriptors, LBP features are used to extract spatial information in each selected frame. Next, multiscale DCT is performed along the ordinate axis of the obtained LBP features to extract spatial information. Representing both spatial and temporal information, the obtained high-level descriptors (LBP-MDCT features) are finally fed into a SVM (Support Vector Machine) classifier to determine whether the input video is a facial attack or valid access. Compared with state of the art, the excellent experimental results of the proposed method on two benchmarking datasets (Replay-Attack and CASIA-FASD dataset) have demonstrated its effectiveness.

Keywords: Video-based face spoofing · LBP · Multiscale DCT · Replay-attack · CASIA-FASD

1 Introduction

Spoofing attack is a kind of presentation attack which is targeted towards fooling a biometric system into recognizing an illegitimate user as a genuine one by means of presenting to the sensor a synthetic forged version of the original biometric trait [1]. Currently it is an accepted fact that, many biometric modalities are vulnerable to spoofing attacks [2–8] wherein face spoofing is of paramount importance to be dealt with. For one thing, face is one of the biometric traits with the highest potential impact both from an economic and a social point of view [1]. For another thing, compared with other spoofing attacks, face spoofing is very easy to conduct due to its low-cost and low-tech features. Facial traits are widely available on the Internet including personal websites and social networks. Moreover, an imposter can easily collect photographs or videos from a genuine user at distance.

Y.Q. Shi et al. (Eds.): IWDW 2016, LNCS 10082, pp. 16–28, 2017.
DOI: 10.1007/978-3-319-53465-7_2

Generally speaking, face spoofing attacks can be divided into three categories: photo attacks, video attacks and mask attacks. In the case of a photo attack, the spoofing artefact is a photograph of the genuine client which can be presented to the sensor on a paper or the screen of a digital device. Video attacks are also referred to as replay attacks since they are carried out by replaying a video of the valid user. A mask attack is a fraudulent access attempt using a 3D mask of the legitimate user enrolled in the system.

With the release of several public face spoofing databases [9–12], many publications addressing face spoofing problems have appeared in the last few years. Overlooking methods dealing with mask attacks which are beyond the scope of this paper, the majority of anti-spoofing methods can be classified into two parts, namely photo-based methods and video-based methods. According to [13], photo-based approaches are not suitable for directly detecting video attacks, especially for high resolution videos. It is more difficult to detect video attacks since the dynamics of the video makes the biometric data more realistic. Besides, video attacks tend to have less degradations and fewer artefacts generated during quantization and discretization. Heretofore, while a number of photo-based schemes have been proposed in the literature, the development of research regarding video-based schemes is not up to par.

Motivated by the strong energy compaction property, DCT has been used extensively in image processing tasks. However, it is adopted to extract static information by performing on a frame-by-frame basis in the literature. Heretofore, it has not beed used for dynamic information extraction. In this paper, we propose an innovative way of applying multiscale DCT to LBP features to represent spatial-temporal information. To the best of our knowledge, this is the first attempt of processing video-based face spoofing attacks by performing multiscale DCT on LBP features in the aim of extracting facial dynamic information. While the proposed scheme is very simple and time-saving, our experiments on two benchmarking datasets showed excellent performance which indicated its effectiveness.

The remainder of the paper is organized as follows. Section 2 briefly reviews the existing methods for detecting face spoofing attacks. Section 3 demonstrates the proposed approach, followed by experimental results in Sect. 4. Finally, conclusions and future work are given in Sect. 5.

2 Related Work

Depending on the type of information used, the majority of existing solutions for face spoofing consists of two parts: solutions using spatial information, solutions using both spatial and temporal information. In order to extract spatial information, the first type of solutions are mainly based on the analysis of face texture using different image processing methods. One of the first attempts towards face spoofing was made in [14], where Fourier spectra of a single face image or face image sequences was analysed. Later, Difference of Gaussian (DoG) was adopted to explore specific frequency information in [10,15]. In [16], Gabor

wavelet was used to enhance texture representation while Histogram Oriented Gradients (HOG) was utilized to introduce local shape description. Proven as an effective tool, Local Binary Pattern (LBP) and its extensions were employed in many works such as [11,13,16–18]. Recently, the authors of [19] created a method for identifying computer-generated facial image which is based on smoothness property of the faces presented by edges and human skin's characteristic via local entropy.

Designed to take advantage of both spatial and temporal information, the second type of anti-spoofing schemes usually achieve very competitive performance, at a cost, in some cases, more time-consuming. An example of such methods is the work in [20] which introduced Local Binary Patterns from Three Orthogonal Planes (LBP-TOP) to combine both space and time information into a single multiresolution texture descriptor. Inspired by the special property of Dynamic Mode Decomposition (DMD), a pipeline of DMD + LBP + SVM was proposed for countering spoof attacks in face anti-spoofing [21]. In [22], the authors combined Multiscale Binarized Statistical Image Features on Three Orthogonal Planes (MBSIF-TOP) and Multiscale Local Phase Quantization Representation on Three Orthogonal Planes (MLPQ-TOP) in the aim of improving the robustness of spoofing attack detector. Recently, a low-level feature descriptor was formed by extracting time-spectral information from the video [23]. Except these texture-based methods, spatial and temporal information are integrated from some other perspectives such as analyzing noise signatures generated by the recapturing process [24].

While the above schemes exploit spatial-temporal information and improve the detection performance, they tend to be more complex and time-consuming. In comparison, the proposed scheme is simple and time-saving. For one thing, the implementations of both LBP and DCT are of high efficiency. For another thing, the proposed method only requires a few of frames instead of the entire frames contained in a video. Furthermore, owing to the strong energy compaction property of DCT, only one or several DCT components are required to form the final descriptor. The scheme as a whole is thus easy to implement and efficient. As for the indispensable classifier in face anti-spoofing task, we use SVM as a back-end classification engine. Our experimental results, strictly following the published experimental protocols, showed that the proposed approach achieved the best performance reported so far. Indeed, we attained zero HTER, i.e., perfect classification, on both the development and test set of Replay-Attack dataset.

3 Proposed Method

In this section, we present an algorithm for video-based attempted spoofing attack detection. The algorithm comprises four main steps: face extraction, low-level descriptor extraction, high-level descriptor extraction, and classification. The overall pipeline is shown in Fig. 1. First, we extract face images from a target video at a fixed time interval. Second, the low-level descriptors, i.e., LBP features are generated for each extracted face image. After that, we perform

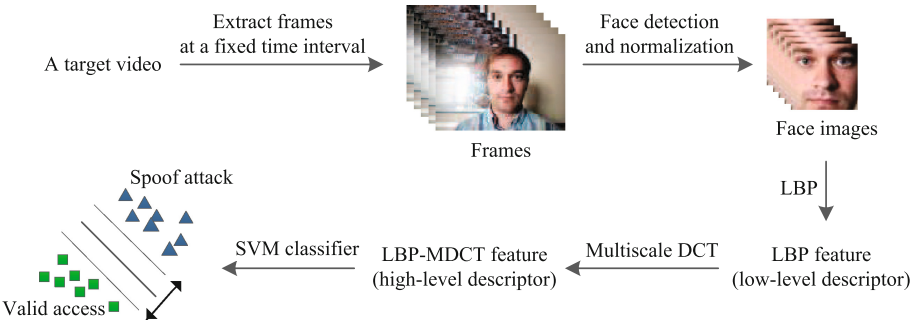


Fig. 1. Flow chart of the proposed method

multiscale DCT on the low-level descriptors to obtain the high-level descriptors (LBP-MDCT features). Finally, the high-level descriptors are fed into a SVM classifier to determine whether the target video is a spoof attack or valid access.

3.1 Face Extraction

Instead of the entire images contained in an input video, only a few of frames are utilized in this work for the sake of efficiency. To this end, we extract frames at a fixed time interval I for each input video. For example, if $I = 1$, the first frame, the third frame, the fifth frame et al. are extracted; if $I = 2$, the first frame, the fourth frame, the seventh frame et al. are extracted. Using the Viola-Jones face detection algorithm [25], we detect face region of each selected frame and then geometrically normalize the face region into 64×64 pixels. Given a video containing M frames, we attain $\lceil \frac{M}{I+1} \rceil$ normalized face images at the end of this stage.

3.2 Low-Level Descriptor Extraction

Local Binary Patterns (LBP), a simple yet a powerful gray-scale invariant texture representation, extracts spatial information based on comparisons of gray values between each pixel and its neighbours. Since it was proposed, many variants have been developed. In this paper, we use the uniform LBP (LBP^{u2}) which only considers labels containing at most two 0−1 or 1−0 transitions. Conventionally, uniform LBP is denoted as $\text{LBP}^{u2}_{P,R}$, where P and R stand for the number of used neighbourhood pixels and radius respectively. Assuming $N = \lceil \frac{M}{I+1} \rceil$ face images are extracted in total, after we obtain a LBP feature vector of length 59, namely a simple normalized histogram of $\text{LBP}^{u2}_{8,1}$ codes for each selected face, all the feature vectors are parallelized, generating a LBP feature matrix of size $N \times 59$ for each video which is adopted as a low-level descriptor in this work.

3.3 High-Level Descriptor Extraction

In this phase, DCT operation is applied to the LBP feature matrix in the aim of extracting temporal information. Figure 2 demonstrates the generation of a LBP-DCT feature. Specifically, after the formation of the $N \times 59$ LBP feature matrix, we perform DCT along its ordinate axis, i.e., the time axis of the entire video, generating many DCT components for each column. Given an input $f(n)$, its 1D-DCT transform can be represented as:

$$F(k) = \alpha(k) \sum_{n=0}^{N-1} f(n) \cos[\frac{(2n+1)k\pi}{2N}] \,, \tag{1}$$

where $0 \leqslant k \leqslant n-1$. Since the strong energy compaction property of DCT, the direct component (DC) concentrates the majority of energy. Taking advantage of this special property, there is no need to utilize all the DCT components. With this in mind, only C components are used in this stage. Specifically, in the case of $C = 1$, only the direct component is extracted; in the case of $C = 2$, the direct component along with the first alternating component are extracted. As a result of this process, we end up with a $59 \times C$ LBP-DCT feature matrix. Of particular worth to mention here is the benefits brought by neglecting most of the DCT components which achieves dimension reduction. Not only does dimension reduction decreases computational complexity, it also improves detection efficiency.

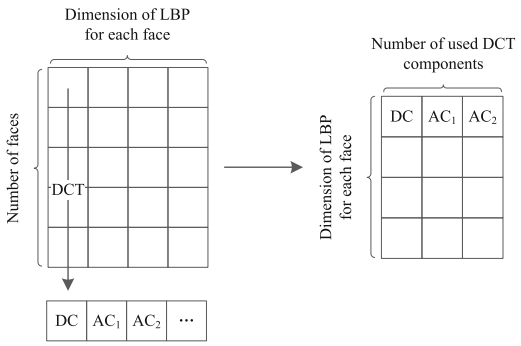


Fig. 2. Demonstration of high-level descriptor extraction

As demonstrated in many previous work [26–28], applying multiscale feature is instrumental in improving performance. Therefore, in order to achieve a better performance, here we perform 3-scale DCT on LBP matrices. In the first scale, we calculate DCT on an entire LBP matrix, generating a $59 \times C$ LBP-DCT feature matrix. In the second scale, the entire LBP matrix is divided into two parts: the first and second part consisting of LBP vectors obtained from the first $\lfloor \frac{N}{2} \rfloor$ frames and the following $\lfloor \frac{N}{2} \rfloor$ frames, respectively. Subsequently,

DCT operation is performed on each part separately, resulting in two LBP-DCT features. Similarly, we obtain four LBP-DCT features in the third scale by partitioning the LBP matrix into four parts. Hence, we end up with seven LBP-DCT features in total through this procedure. Last but not least, the resultant seven LBP-DCT features are concatenated to form a final matrix which is referred to as a high-level descriptor (LBP-MDCT feature) in the present paper.

3.4 Classification

The last step of the proposed algorithm is to use a discriminative machine learning approach to find an appropriate classification model in order to decide whether an input video is an attempted attack or not. In this paper, we choose Support Vector Machine (SVM) [29] algorithm with Radial Basis Function (RBF) kernel. After we feed the high-level descriptors into a SVM, the detection of face spoofing attacks is completed according to the output of the SVM.

4 Experimental Results

In this section, the experimental results of the proposed method are presented and discussed. Prior to proceeding with the experiments, we give an overview of the datasets and experimental protocols employed to evaluate performance of different countermeasures.

4.1 Datasets

In this work, we consider two benchmarking datasets.

Replay-Attack Dataset [11]: This dataset consists of short video recordings of about 10 s of both real access and spoofing attacks to a face recognition system. It contains 1200 videos (200 real access and 1000 spoof attacks) of 50 identities and the samples are taken in 3 different scenarios with 2 different illumination and support conditions.

CASIA-FASD Dataset [10]: This dataset comprises 600 videos (150 real access and 450 attacks) of 50 subjects. It introduces face attacks with a varying degree of imaging quality, namely low quality (captured using an old USB camera with a resolution of 640×480 pixels), normal quality (captured using a new USB camera with a resolution of 480×640 pixels) and high quality (captured using a Sony NEX-5 with a resolution of 1920×1080 pixels). It also considers three different types of attacks including warped photo attacks, cut photo attacks and video attacks.

4.2 Experimental Protocols

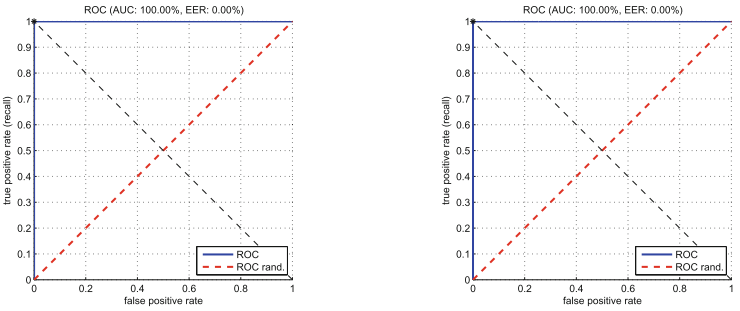
In order to fairly measure performance of different face spoofing detection systems, we adopt the widely-used Half Total Error Rate (HTER) as the evaluation parameter. As the following equation shows, HTER is half of the sum of the

False Acceptance Rate (FAR) and the False Rejection Rate (FRR). Since both the FAR and the FRR depend on a threshold τ, increasing one usually decreases the other. For this reason, HTER is conventionally calculated in a specific operating point of the Receiver Operating Characteristic (ROC) curve in which the FAR is equal to the FRR, known as the Equal Error Rate (ERR). Apparently, the lower the HTER value is, the better the approach is.

$$\text{HTER} = \frac{1}{2}(\text{FAR} + \text{FRR}). \tag{2}$$

Protocol I: In this evaluation protocol, we use the Replay-Attack dataset. This dataset is divided into three partitions: training, development and testing set. The training set is used for training the classifier itself and the development set is used for choosing the threshold τ on the EER. Finally, the testing set is used to report the HTER value.

Protocol II: Herein, we use the CASIA-FASD dataset which is divided into two disjoint subsets for training and testing (240 and 360 videos, respectively). The training subset is used to build a classifier and the testing set is used to report the HTER value.



(a) (b)

Fig. 3. ROC curves of Replay-Attack dataset when $I = 72$ & $C = 1$. (a) Development set (b) Test set

4.3 Effectiveness of the Proposed Scheme on Replay-Attack Dataset

Table 1 shows the performance of this countermeasure on Replay-Attack dataset in terms of HTER with different numbers of used DCT components C and different values of interval I. Note that, there is a limitation for interval I under a specific C. For instance, given $C = 1$, at least 4 frames are required as we perform 3-scale DCT on LBP features. Consequently, when $C = 1$, the maximum value of interval I is 72 due to the minimum number of frames contained in each video in this dataset is 221. Subject to limited space, only part of the parameter sittings are shown in Table 1.

Table 1. Results in HTER (%) of the proposed method on Replay-Attack and CASIA-FASD dataset. Here, C: number of used DCT components, I: interval for extracting frames, Dev: on development set and Test: on test set

C	I	Replay-Attack		CASIA-FASD
		Dev	Test	Test
1	1	0.00	0.00	20.00
	2	0.00	0.00	19.07
	3	0.00	0.00	20.00
	4	0.00	0.00	18.89
2	1	0.00	0.00	20.00
	2	0.00	0.00	19.07
	3	0.00	0.00	20.00
	4	0.00	0.00	19.26
3	1	0.00	0.00	18.89
	2	0.00	0.00	19.26
	3	0.00	0.00	19.07
	4	0.00	0.00	18.43
4	1	0.00	0.00	18.89
	2	0.00	0.00	20.00
	3	0.00	0.00	18.25
	4	0.00	0.00	18.06
5	1	0.00	0.00	18.89
	2	0.00	0.00	19.18
	3	0.00	0.00	19.18
	4	0.00	0.00	18.89

Surprisingly, all the parameter settings lead to HTER values of 0 on both development set and testing set of Replay-Attack dataset. In other words, all the parameter settings achieved perfect classification. Apparently, the greater I is, the fewer frames are involved, thus the faster the detection is and the lower the computational complexity is. Likewise, the less C is, the faster and the simpler the detection is. Therefore, we consider $I = 72$ & $C = 1$ as the best setting for the sake of efficiency and simplicity. Figure 3 shows the corresponding ROC curves when $I = 72$ & $C = 1$.

4.4 Effectiveness of the Proposed Scheme on CASIA-FASD Dataset

Table 1 also shows the HTER of this countermeasure on CASIA-FASD dataset with different numbers of used DCT components C and different interval values I. Same as Replay-Attack dataset, subject to the minimum number of frames

contained in each video in CASIA-FASD dataset, there is a maximum value for interval I given a specific C. We also omit some of the results due to limited space.

From Table 1 we notice that while HTER value of the proposed countermeasure on this dataset fluctuates around 19.00%, $I = 4$ & $C = 4$ gives an HTER of 18.06s%, recording the best result on the testing set of CASIA-FASD dataset. Therefore, we consider $I = 4$ & $C = 4$ as the best setting for CASIA-FASD dataset. Figure 4 shows the corresponding ROC curve when $I = 4$ & $C = 4$. As for the relationship between the parameter setting (I and C) and performance fluctuation, it falls into one direction of our future work. Unlike perfect classification attained on Replay-Attack dataset, performance on this dataset is less satisfactory. We believe this is because CASIA-FASD dataset is more challenging than the former dataset. For instance, it incorporates cut photo attack which is a kind of printed 2D attacks. As it simulates eye blinking with perforated eyes, this new kind of attack is difficult to detect. In addition, CASIA-FASD dataset is more heterogeneous than Replay-Attack dataset as it contains three types of attacks (warped photo attack, cut photo attack and reply attack) captured in three different settings (low, normal and high resolutions).

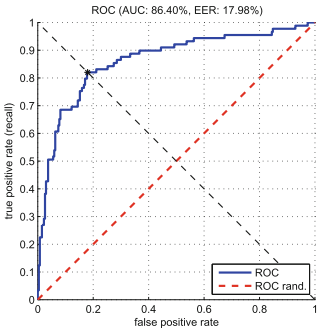


Fig. 4. ROC curve of CASIA dataset when $I = 4$ & $C = 4$

4.5 Comparisons with State of the Art

Finally, Table 2 shows a comparison among state-of-the-art methods. Consistent with our observation that CASIA-FASD dataset is more challenging than Replay-Attack dataset, all the methods show a better performance on Replay-Attack dataset than CASIA-FASD dataset.

We can see from Table 2 that the proposed approach achieves the best performance among other approaches. For Replay-Attack dataset, it recorded zero HTER, i.e., perfect classification. For CASIA-FASD dataset, it recorded 18.06% HTER which also outperforms all the other methods with a minimum superiority of 3.69%. Using LBP for feature extraction, Chingovska et al.'s methods [11] show moderate discriminability with about 15.00% HTER. The best performance

Table 2. Comparison of HTER (%) on test sets for the proposed method with respect to state-of-the-art methods.

	Algorithm	Replay-Attack	CASIA-FASD
Chingovska et al. (2012) [11]	LBP + LDA	13.87	-
	LBP + SVM	18.17	-
Pereira et al. (2013) [30]	Motion Correlation	11.79	30.33
	LBP	15.45	23.19
	LBP-TOP	8.51	23.75
Tirunagari et al.(2015) [21]	DMD + LBP + SVM^E	0.00	-
	DMD + LBP + SVM^F	3.75	21.75
	DMD + SVM^F	7.50	29.50
	PCA + SVM^F	21.50	33.50
	PCA + LBP + SVM^F	17.11	24.50
	PCA + LBP + SVM^E	20.50	-
Proposed method	LBP + MDCT + SVM	0.00	18.06

Here, E: on entire frames and F: on face regions

achieved in Pereira et al.'s work [30] is based on LBP-TOP and the corresponding HTER is 8.51% on Replay-Attack dataset. Although Tirunagari et al.'s method DMD + LBP + SVM^E [21] also gives an HTER of 0% on Replay-Attack dataset, it requires 240 frames from each video while the proposed pipeline only requires 4 frames. As face spoofing detection is aimed for real-time application, the fewer frames the approach uses, the less time the detection takes, thus, the higher the efficiency is and the better the approach is. Therefore, the proposed approach is better than approach DMD + LBP + SVM^E. More importantly, instead of the face regions, pipeline DMD + LBP + SVM^E requires the entire video frames. Once only the face regions are given (DMD + LBP + SVM^F), the performance degrades to 3.75% HTER. This degradation is made by the fact that the backgrounds of videos for different types of attacks in CASIA-FASD dataset are different which may well benefit the detection performance. However, in real life, this difference tends to be non-existent for a certain biometric system. From this point of view, the perfect classification attained by DMD + LBP + SVM^E is meaningless. With all these in mind, the proposed pipeline is superior to all the others including Tirunagari et al.'s method.

In addition to the excellent performance, our approach is of low complexity and high efficiency as it only requires a few of frames rather than the entire frames contained in a video. Besides, the implementation of both LBP and DCT operations are time-saving. For videos from Replay-Attack or CASIA-FASD dataset, it only takes about 0.12 s to obtain a LBP vector for each frame. Given a LBP matrix of a video from the two datasets, the calculation of a LBP-MDCT feature only costs about 0.02 s whatever the parameter sitting is. Moreover, taking advantage of the strong energy compaction property of DCT, only one or several DCT components are used in the stage of high-level descriptor extraction which

realises dimension reduction. The scheme as a whole is thus easy to implement and efficient. With low complexity and high efficiency, the proposed method is very competitive and promising for face anti-spoofing which is aimed for real life scenario.

5 Conclusions and Future Work

The present paper addressed the problem of video-based face spoofing detection using LBP and multiscale DCT. $\text{LBP}_{8,1}^{u2}$ features are adopted as the low-level descriptors to exact spatial information while multiscale DCT are implemented along the ordinate axis of LBP features in order to take advantage of temporal information at the same time. Therefore, the resultant LBP-MDCT features represents spatial-temporal information. Excellent experimental results demonstrated that multiscale DCT can indeed capture facial dynamic information which makes a significant contribution to distinguish spoof attacks from valid access.

The proposed method was comprehensively evaluated on two benchmarking datasets using standard protocols and shown to be superior to other approaches reported to date. For Replay-Attack dataset, it recorded zero HTER, i.e., perfect classification. For CASIA-FASD dataset, it recorded 18.06% HTER which also outperforms all the other methods with a minimum superiority of 3.69%. We attribute this performance superiority to (1) the capacity of LBP to extract texture information on a frame-by-frame basis, (2) the capacity of multiscale DCT to extract temporal dynamics of video sequences and (3) the choice of pipeline within which multiscale DCT is deployed along the ordinate axis of LBP matrices to capture spatial-temporal information. Note that, instead of the strength of LBP or DCT alone, it is the unique combination of the two descriptors that lead to the superiority of the proposed method. In addition to the excellent performance, our approach is time-saving and very simple to conduct. Taking its outstanding performance, low complexity and high efficiency into account, the proposed method is very competitive and promising for practical application of face anti-spoofing.

Our focus of future research will include a further investigation into the relationship between the skipped frames and detection performance on CASIA-FASD dataset. It would also be interesting to exploit other types of LBP descriptor with the same pipeline and compare the performance. Another future research direction relates to replacing LBP or DCT with other effective operations in the context of using spatial-temporal information.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (61272414) and the research funding of State Key Laboratory of Information Security (2016-MS-07).

References

1. Galbally, J., Marcel, S., Fierrez, J.: Biometric antispoofing methods: a survey in face recognition. IEEE Access **2**, 1530–1552 (2014)
2. Anjos, A., Marcel, S.: Counter-measures to photo attacks in face recognition: a public database and a baseline. In: 2011 IEEE International Joint Conference on Biometrics (IJCB), pp. 1–7. IEEE Press, Washington, DC (2011)
3. Galbally, J., Fierrez, J., Alonso-Fernandez, F., Martinez-Diaz, M.: Evaluation of direct attacks to fingerprint verification systems. Telecommun. Syst. **47**(3–4), 243–254 (2011)
4. Mjaaland, B.B., Bours, P., Gligoroski, D.: Walk the walk: attacking gait biometrics by imitation. In: Burmester, M., Tsudik, G., Magliveras, S., Ilić, I. (eds.) ISC 2010. LNCS, vol. 6531, pp. 361–380. Springer, Heidelberg (2011). doi:10.1007/978-3-642-18178-8_31
5. Chen, H., Valizadegan, H., Jackson, C., Soltysiak, S., Jain, A.K.: Fake hands: spoofing hand geometry systems. In: 2005 Biometrics Consortium Conference (BCC) (2005)
6. Bin, Q., Jian-Fei, P., Guang-Zhong, C., Ge-Guo, D.: The anti-spoofing study of vein identification system. In: International Conference on Computational Intelligence and Security (ICCIS), pp. 357–360 (2009)
7. Akhtar, Z., Fumera, G., Marcialis, G.L., Roli, F.: Evaluation of serial and parallel multibiometric systems under spoofing attacks. In: 5th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 283–288 (2012)
8. Tome, P., Vanoni, M., Marcel, S.: On the vulnerability of finger vein recognition to spoofing. In: International Conference of the Biometrics Special Interest Group (BIOSIG), pp. 1–10 (2014)
9. Tan, X., Li, Y., Liu, J., Jiang, L.: Face liveness detection from a single image with sparse low rank bilinear discriminative model. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6316, pp. 504–517. Springer, Heidelberg (2010). doi:10.1007/978-3-642-15567-3_37
10. Zhang, Z., Yan, J., Liu, S., Lei, Z., Yi, D., Li, S.: A face antispoofing database with diverse attacks. In: 2012 5th IAPR International Conference on Biometrics (ICB), pp. 26–31. IEEE Press (2012)
11. Chingovska, I., Anjos, A., Marcel, S.: On the effectiveness of local binary patterns in face anti-spoofing. In: International Conference of Biometrics Special Interest Group (BIOSIG), pp. 1–7. IEEE Press, Darmstadt (2012)
12. Erdogmus, N., Marcel, S.: Spoofing in 2D face recognition with 3D masks and anti-spoofing with kinect. In: 6th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–6 (2013)
13. Pinto, A., Schwartz, W.R., Pedrini, H., de Rezende Rocha, A.: Using visual rhythms for detecting video-based facial spoof attacks. IEEE Trans. Inf. Forensics Secur. **10**(5), 1025–1038 (2015)
14. Li, J., Wang, Y., Tan, T., Jain, A.K.: Live face detection based on the analysis of fourier spectra. Proc. SPIE **5404**, 296–303 (2004)
15. Peixoto, B., Michelassi, C., Rocha, A.: Face liveness detection under bad illumination conditions. In: 2011 18th IEEE International Conference on Image Processing (ICIP), pp. 3557–3560. IEEE Press, Brussels (2011)
16. Maatta, J., Hadid, A., Pietikäinen, M.: Face spoofing detection from single images using texture and local shape analysis. IET Biometrics **1**(1), 3–10 (2012)

17. Kose, N., Dugelay, J.L.: Classification of captured and recaptured images to detect photograph spoofing. In: 2012 International Conference on Informatics, Electronics and Vision (ICIEV), Dhaka, pp. 1027–1032 (2012)
18. Maatta, J., Hadid, A., Pietikäinen, M.: Face spoofing detection from single images using micro-texture analysis. In: 2011 International Joint Conference on Biometrics (IJCB), pp. 1–7. IEEE Press, Washington, DC (2011)
19. Nguyen, H.H., Nguyen-Son, H.-Q., Nguyen, T.D., Echizen, I.: Discriminating between computer-generated facial images and natural ones using smoothness property and local entropy. In: Shi, Y.-Q., Kim, H.J., Pérez-González, F., Echizen, I. (eds.) IWDW 2015. LNCS, vol. 9569, pp. 39–50. Springer, Heidelberg (2016). doi:10.1007/978-3-319-31960-5_4
20. de Freitas Pereira, T., Anjos, A., De Martino, J.M., Marcel, S.: *LBP–TOP* based countermeasure against face spoofing attacks. In: Park, J.-I., Kim, J. (eds.) ACCV 2012. LNCS, vol. 7728, pp. 121–132. Springer, Heidelberg (2013). doi:10.1007/978-3-642-37410-4_11
21. Tirunagari, S., Poh, N., Windridge, D., Iorliam, A., Suki, N., Ho, A.T.S.: Detection of face spoofing using visual dynamics. IEEE Trans. Inf. Forensics Secur. **10**(4), 762–777 (2015)
22. Arashloo, S.R., Kittler, J., Christmas, W.: Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features. IEEE Trans. Inf. Forensics Secur. **10**(11), 2396–2407 (2015)
23. Pinto, A., Pedrini, H., Schwartz, W.R., Rocha, A.: Face spoofing detection through visual codebooks of spectral temporal cubes. IEEE Trans. Image Process. **24**(12), 4726–4740 (2015)
24. da Silva Pinto, A., Pedrini, H., Schwartz, W., Rocha, A.: Video-based face spoofing detection through visual rhythm analysis. In: 2012 25th SIBGRAPI Conference on Graphics, Patterns and Images, Ouro Preto, pp. 221–228 (2012)
25. Viola, P., Jones, M.J.: Robust real-time face detection. Int. J. Comput. Vis. **57**(2), 137–154 (2004)
26. Arashloo, S.R., Kittler, J.: Class-specific kernel fusion of multiple descriptors for face verification using multiscale binarised statistical image features. IEEE Trans. Inf. Forensics Secur. **9**(12), 2100–2109 (2014)
27. Arashloo, S.R., Kittler, J.: Dynamic texture recognition using multiscale binarized statistical image features. IEEE Trans. Multimedia **16**(8), 2099–2109 (2014)
28. Chan, C.H., Tahir, M.A., Kittler, J., Pietikainen, M.: Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors. IEEE Trans. Pattern Anal. Mach. Intell. **35**(5), 1164–1177 (2013)
29. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
30. de Freitas Pereira, T., Komulainen, J., Anjos, A., De Martino, J.M., Hadid, A., Pietikäinen, M., Marcel, S.: Face liveness detection using dynamic texture. EURASIP J. Image Video Process. **2014**(2), 1–15 (2014)

http://www.springer.com/978-3-319-53464-0

Digital Forensics and Watermarking
15th International Workshop, IWDW 2016, Beijing, China, September 17-19, 2016, Revised Selected Papers
Shi, Y.Q.; Kim, H.J.; Perez-Gonzalez, F.; Liu, F. (Eds.)
2017, XIV, 618 p. 269 illus., Softcover
ISBN: 978-3-319-53464-0