Proceedings of the 2014 International Conference on Wavelet Analysis and Pattern Recognition, Lanzhou, 13-16 July, 2014

FACE LIVENESS DETECTION USING ANALYSIS OF FOURIER SPECTRA BASED ON HAIR

WEIWEN LIU

School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China
EMAIL: liu.weiwen@mail.scut.edu.cn

Abstract:

The existing face liveness detection algorithms have satisfying performance on spoofing photos by detecting the biological motion, including eye blinking, head rotation and mouth movement. However, the biological motions can be simulated by using video, which is called a video playback attack. This paper proposes a countermeasure to a video playback attack by analyzing the detail of hair using the Fourier spectra. We apply a flashlight to enhance the difference between the real person and the video playback attack. Experimental results suggest that the proposed method has an encouraging performance on face liveness detection.

Keywords:

Face liveness detection; Hair extraction; Fourier spectra

1. Introduction

The accuracy of face recognition has been improved significantly due to the improvement on the hardware and the technique in computer vision. Face recognition becomes one of the most efficient methods of biometric identification because of its non-intrusive interaction. Since it has been successfully applied in many security applications, i.e. the identity authentication system [23], potential criminals searching systems [24] etc. Therefore, the incentive of defeating face recognition systems increases.

An adversary intentionally misleads the decision of the system in order to get the authorization by manipulating input samples. For example, printed photographs [1] and printed photographs with perforated eyes regions [2] are common used methods in which a photo of a real user is presented to the camera to mislead the system. The existed defense methods to the static image have been well studied and achieved satisfying results [3]. Recently, more advanced evasion attacks play videos by tablets or smart phones [4] to increase the difficulty of the face liveness detection. Spoofing videos have more physiological clues than photos since videos can express the signs of life, i.e. facial expression, eye blinking or lips movement. High-quality spoofing video is almost as clear as a real face. Thus, the development of the face liveness detection system is necessary to improve the security of the system.

The existing counter systems for 2D spoofing attacks can be classified in two categories: the user-cooperative systems [5-9] and the non-user-cooperative systems [10-18]. The user-cooperative systems require users to do specific tasks, e.g. slowly move head [9] or speak a word [6]. This method is inconvenient to users, and even worse, it is ineffective if users refuse to cooperate. The non-user-cooperative systems do not require users to follow an instruction. For example, the texture-based methods explore the artifacts' texture and special effects appearing when using the non-skin texture tools to display the images [12-14]; the motion-based methods explore the unnatural movements produced by spoofing attackers who use image sequences [16]; the liveness-based methods detect the evidence indicator of liveness produced by real persons [15]. Some methods require auxiliary devices in classifying the liveness. Examples of auxiliary devices are infrared equipment [10], 3D cameras and multiple 2D cameras [11]. Although the auxiliary-based methods usually achieve higher accuracy than the one without using auxiliary device, its cost and difficulty of implementation are higher in general.

The quality of a screen which displays a fake image is not as good as a real object. Thus, high frequency components of real face images should be more than a fake face image. Moreover, the standard deviation of frequency components in sequence must be smaller than a real face since the video is composed by a series of photos.

978-1-4799-4211-4/14/$31.00 ©2014 IEEE

As a result, the method based on the analysis of 2D Fourier Spectra, which measures the corresponding high frequency descriptor (HFD) and the frequency dynamics descriptor (FDD), has been proposed [12]. However, this method implicitly assumes that the quality of the screen is low and requires change of pose when detecting. The performance may be downgraded when confronting a clear photo or video.

To improve this method, we propose a method which considers the difference on the hair of a user with and without flashlight. Under the flashlight, the detail of the hair, measured by high frequency components, is shown clearly for a real person in comparison with a fake image or video. The economical auxiliary device, i.e. a flashlight, increases the difference on the hairs of a real and fake object significantly. The advantages of the proposed method are the low computational complexity and implementation cost. The proposed method is evaluated experimentally using the dataset collected by our group.

This paper is organized as follows: section 2 shows the related work of the analysis of Fourier spectra. The proposed face liveness detection method using hair abstraction with extra light is devised in Section 3. Section 4 illustrates experimental results using our collected dataset. The conclusion is given in Section 5

2. Related work

The spoofing attack [3, 6, 7, 17, 25] referring to the illegal access to the face recognition system using fake face samples, including *photograph attack* [3,17], *video playback attack* [6,7] and *mask face attack* [25]. Most researches focus on the 2D spoofing attack [5-18].

Many anti-spoofing detection methods have been proposed for 2D spoofing attack. The degree of synchrony between the lips and the voice extracted from a video sequence based on co-inertia analysis (CoIA) and a fourth based on coupled hidden Markov models (CHMMs) is measured [7]. As a person blinks once every 2 to 4 seconds, the spontaneous eye blinks [17] is detected. These methods are not suitable for video playback attack. The texture of skin is analyzed to distinguish the face displayed on the screen from the real [13, 18]. One of disadvantages of these methods is that it requires restrictive illumination condition. Another kinds of methods require additional devices, for example, the thermal infrared imaging camera [10]. All the methods above need either a complicated computational process or an auxiliary device.

3. Proposed Method

The proposed method is based on the difference of the hair texture with and without the flashlight. Two photos, including ones taken with and without flashlight, are taken from the object. The reason of considering the region of hair because the change of hair under different illumination are more obvious than other parts. The high frequency component (HFD) is calculated for the hair region. The two values of HFDs of the pair of images are applied to classify the input images. The flowchart is shown as follows:

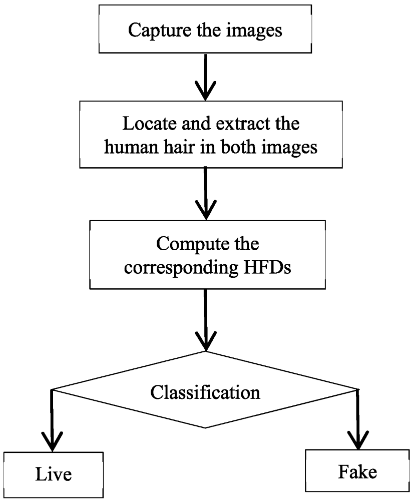
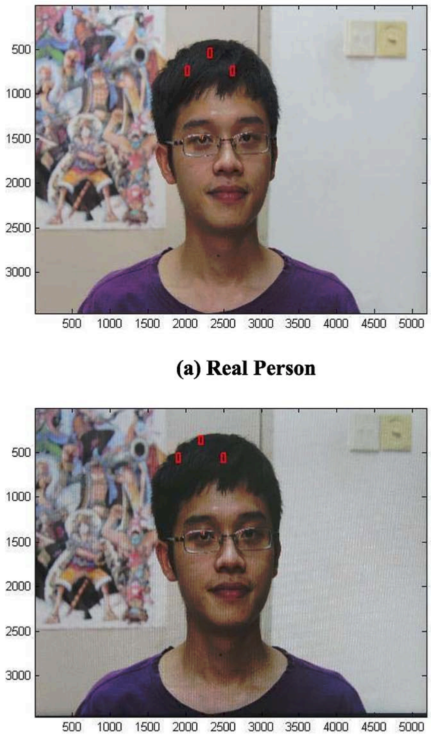

Capture the images
Locate and extract the human hair in both images
Compute the corresponding HFDs
Classification
Live
Fake

Figure 1 Flowchart

3.1 Hair location and extraction

We locate the hair region using a simple model-based face detection algorithm [22]. The algorithm integrate the skin color segmentation and Split up Sparse Network of Winnows (SNoW) [21]. The approach of skin color segmentation selects the skin color areas in order to obtain the face. To extract illumination in sensitive features, the Local Successive Mean Quantization Transform (SMQT) [20] is adopted due to their satisfying results under complicated illumination.

As the images with and without flashlight are taken in short period of time, the head movement is slight. The hair region is determined according to the image without the flashlight and is also applied to the image with flashlight. As the high frequency varies in the hair region, three rectangles (100×50 pixels) are extracted according to the output face positions to calculate the average HFD value to minimize the error caused by the moving by the users. Figure 2(a) and 2(b) display an example of the extracted hair region of a real person and a fake face without flashlight.



(a) Real Person

(b) Fake Face

Figure 2 Extracted hair region without light

3.2 High frequency component

The high frequency descriptor defined in (1) represents certain frequencies of an image. The high frequency should satisfy two conditions: the frequency is greater than a predefined threshold T_{fd} and the value of Fourier transform is greater than another predefined threshold T_f.

$$HFD_i = \frac{\iint_{\Omega=\{(u,v)|\sqrt{u^2+v^2}\}>T_{fd}\cdot f_{max} and |F(u,v)|>T_f} |F(u,v)|dudv}{\{\iint |F(u,v)|dudv - F(0,0)\}} \tag{1}$$

$$\text{HFD} = \frac{HFD_1 + HFD_2 + HFD_3}{3} \tag{2}$$

where HFD_i is the value of high frequency descriptor for rectangle i where i = 1, 2, 3, $F(u,v)$ denotes Fourier transform of an face image, $f_{\max}$ is the highest radius frequency of $F(u,v)$, T_f is a predefined threshold. The denominator represents the total energy in frequency domain which is the sum of Fourier coefficients relative to direct coefficient. Figure 3 illustrates an example of the frequency domain of an input image.

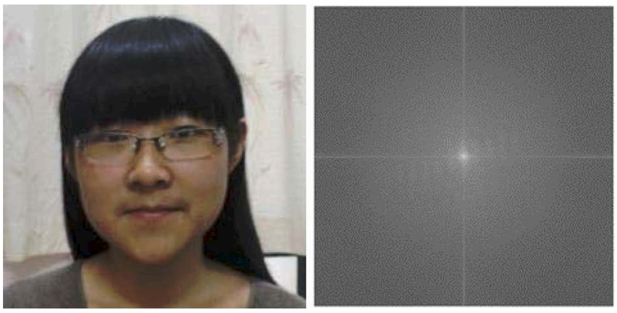
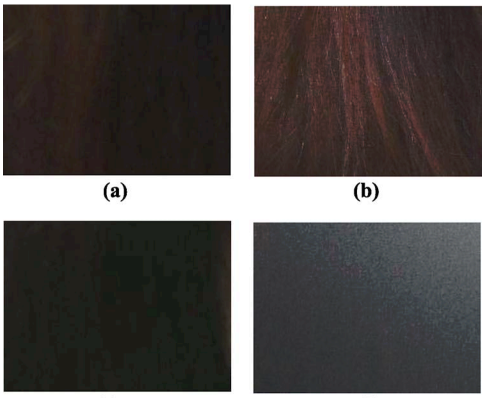


Figure 3 Frequency domain (right) of an input image (left)

3.3 Hair analysis

HFD defined as (2) is calculated for the images of hair with and without flashlight. The value of HFD is applied as features for classification. Figure 4 shows the difference between hair images of live and fake face. Figure (a) and (c), which are taken without flashlight, are both rather vague. However, figure (b) and (d) show the images taken with flashlight. Figure (b) is for real face, and it display the texture of hair. Figure (d) is for fake face, and it is even vaguer. As a result, the flashlight increases the difference between the hairs of real and fake person.



(a) **(b)**

(c) **(d)**

Figure 4 Difference between the hair of the real and fake object: (a) hair of a live person without light; (b) hair of a live person with light; (c) hair of a fake person without light; (d) hair of a fake person with light.

4. Experiments

This section provides an analysis of Fourier spectra using different illumination conditions based on hair. The experiment is carried out using the dataset descripted in Section 4.1. Tfd=2/3, Tf=1.0 are selected for our method. The normal densities based quadratic classifier, the BP Neural Network, Bayes Classifier, Radial Basis Function (RBF) Network and Decision Tree are applied as classifiers. Our method is compared with the Ordinary Fourier Spectra Analysis method (OFSA) [12] 5-fold cross validation is used.

4.1 Dataset Collection

We constructed a photo imposter database using a Canon EOS 600D DSLR camera. The images in the database are collected in three different indoor rooms within one day. The illumination settings of the three places are the same. The distance between the camera and the subject is 3 meters. Totally 21 person were invited to participate in this work. Four photos are taken for each person, including live face without light, live face with light, fake face without light and fake face with light. The size of each photo is 5184×3456 pixel in JPG format. The fake face is contrasted by displaying the photos using a graphoscope. Then we re-photographed the images using the same camera. As a result, 84 photos, including half of them are legitimate and malicious samples, are contained in the dataset.

Since the hair condition is severely restricted based on the algorithm, only the photos with a person with fringes over the forehead are considered. In this experiment, we selected 11 out of 21 groups in the experiment.

4.2. Result and Discussion

Table 1 Example HFD values of samples

Sample	Real face without light	Real face with light	Fake face without light	Fake face with light
1	4.1399	4.1922	2.6983	2.4378
2	3.8892	4.3575	4.2077	2.4899
3	4.6498	5.0953	2.8807	2.8107
4	4.4516	5.5986	3.6489	3.0788

The values of HFD of the samples are illustrated in Table 1. The result suggests that the flashlight make the detail of the hairs of a real person, but not fake one, clearer. As the screen is much smooth than a real face, the reflection is more intense for a screen. As a result, some details are lost for fake objects.

Table 2 The error rate using different classifiers

Classifier	Accuracy
Quadratic Classifier	100.00%
BP Neural Network	16.67%
Bayes Classifier	0.00%
RBF Network	16.67%
Decision Tree	0.00%

The accuracies of the classifiers are reported in Table 2. The normal densities based quadratic classifier, naïve Bayes classifier and the decision tree can all achieve 100% accuracy. The linear classifier, the BP neural network and RBF function achieve 83.3% accuracy. Thus, the simplest and the most computationally efficient classifier, i.e. Bayes classifier, should be selected when implementing the anti-spoofing solutions in practice.

Table 3 Results of comparison of our method and the ordinary Fourier spectra analysis method (OFSA)

	Accuracy	Precision	Recall
OFSA	40.48%	44.12%	71.43%
Our method	100%	100%	100%

The comparison result is shown in Table 3. The result indicates that the performance of the proposed method, which is 100% accurate, is more stable and robust in comparison with OFSA, which is only 40.48% accurate.

5. Conclusion and future work

Liveness detection improves the robustness of face recognition system to adversarial attack. This paper proposes a method which considers the difference on the hair of a user with and without flashlight. By using the economical auxiliary device, i.e. a flashlight, the efficiency of the liveness detection has increased. The proposed method is evaluated and compared with the existing method using the dataset containing the users with haircut-fringe hair. The classifiers using the proposed features have satisfying result. The main deficiency of the method is the restriction on the hair cut of the users. Besides, more robust and accurate of the hair extracting programs are needed.

Acknowledgements

This work is supported by a National Natural Science Foundation of China (61272201), and a Fundamental Research Funds for the Central Universities (10561201465).

References

[1] Tan, Xiaoyang, Yi Li, Jun Liu, and Lin Jiang. "Face liveness detection from a single image with sparse low rank bilinear discriminative model." In Computer Vision–ECCV 2010, pp. 504-517. Springer Berlin Heidelberg, 2010.

[2] Zhang, Zhiwei, Junjie Yan, Sifei Liu, Zhen Lei, Dong Yi, and Stan Z. Li. "A face antispoofing database with diverse attacks." In Biometrics (ICB), 2012 5th IAPR International Conference on, pp. 26-31. IEEE, 2012.

[3] Tronci, Roberto, Daniele Muntoni, Gianluca Fadda, Maurizio Pili, Nicola Sirena, Gabriele Murgia, Marco Ristori, and Fabio Roli. "Fusion of multiple clues for photo-attack detection in face recognition systems." In Biometrics (IJCB), 2011 International Joint Conference on, pp. 1-6. IEEE, 2011.

[4] Kahm, O., and Naser Damer. "2D face liveness detection: An overview." InBiometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the, pp. 1-12. IEEE, 2012.

[5] Chetty, Girija. "Biometric liveness checking using multimodal fuzzy fusion." InFuzzy Systems (FUZZ), 2010 IEEE International Conference on, pp. 1-8. IEEE, 2010.

[6] Chetty, Girija, and Michael Wagner. "Biometric person authentication with liveness detection based on audio-visual fusion." International Journal of Biometrics 1, no. 4 (2009): 463-478.

[7] Rúa, Enrique Argones, Hervé Bredin, Carmen García Mateo, Gérard Chollet, and Daniel González Jiménez. "Audio-visual speech asynchrony detection using co-inertia analysis and coupled hidden markov models." Pattern Analysis and Applications 12, no. 3 (2009): 271-284.

[8] Fauve, Benoıt, Hervé Bredin, Walid Karam, Florian Verdet, Aurélien Mayoue, Gérard Chollet, Jean Hennebert et al. "Some results from the biosecure talking face evaluation campaign." In Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, pp. 4137-4140. IEEE, 2008.

[9] Kollreider, Klaus, Hartwig Fronthaler, and Josef Bigun. "Non-intrusive liveness detection by face images." Image and Vision Computing 27, no. 3 (2009): 233-244.

[10] Socolinsky, Diego A., Andrea Selinger, and Joshua D. Neuheisel. "Face recognition with visible and thermal infrared imagery." Computer Vision and Image Understanding 91, no. 1 (2003): 72-114.

[11] Kosmerlj, Marijana, Tom Fladsrud, Erik Hjelmås, and Einar Snekkenes. "Face recognition issues in a border control environment." In Advances in Biometrics, pp. 33-39. Springer Berlin Heidelberg, 2005.

[12] Li, Jiangwei, Yunhong Wang, Tieniu Tan, and Anil K. Jain. "Face liveness detection based on the analysis of fourier spectra." In Defense and Security, pp. 296-303. International Society for Optics and Photonics, 2004.

[13] Chingovska, Ivana, André Anjos, and Sébastien Marcel. "On the effectiveness of local binary patterns in face anti-spoofing." In Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the, pp. 1-7. IEEE, 2012.

[14] Määttä, J., A. Hadid, and M. Pietikäinen. "Face spoofing detection from single images using texture and local shape analysis." IET biometrics 1, no. 1 (2012): 3-10.

[15] Pan, Gang, Zhaohui Wu, and Lin Sun. "Liveness detection for face recognition." Recent advances in face recognition (2008): 109-124.

[16] Bharadwaj, Samarth, Tejas I. Dhamecha, Mayank Vatsa, and Richa Singh. "Computationally Efficient Face Spoofing Detection with Motion Magnification." In Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on, pp. 105-110. IEEE, 2013.

[17] Pan, Gang, Lin Sun, Zhaohui Wu, and Shihong Lao. "Eyeblink-based anti-spoofing in face recognition from a generic webcamera." In Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, pp. 1-8. IEEE, 2007.

[18] Zhang, Zhiwei, Dong Yi, Zhen Lei, and Stan Z. Li. "Face liveness detection by learning multispectral reflectance distributions." In Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on, pp. 436-441. IEEE, 2011.

[19] Basri, Ronen, and David W. Jacobs. "Lambertian reflectance and linear subspaces." Pattern Analysis and Machine Intelligence, IEEE Transactions on25, no. 2 (2003): 218-233.

[20] Nilsson, Mikael, Mattias Dahl, and Ingvar Claesson. "The successive mean quantization transform." In Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP'05). IEEE International Conference on, vol. 4, pp. iv-429. IEEE, 2005.

[21] Nilsson, Mikael, Jörgen Nordberg, and Ingvar Claesson. "Face detection using local SMQT features and split up snow classifier." In Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on, vol. 2, pp. II-589. IEEE, 2007.

[22] Hsu, Rein-Lien, Mohamed Abdel-Mottaleb, and Anil K. Jain. "Face detection in color images." Pattern Analysis and Machine Intelligence, IEEE Transactions on 24, no. 5 (2002): 696-706.

[23] Szwoch, Mariusz, and Paweł Pieniążek. "Eye blink based detection of liveness in biometric authentication systems using conditional random fields." Computer Vision and Graphics. Springer Berlin Heidelberg (2012):669-676.

[24] Singh, Manminder, and A. S. Arora. "Face Recognition and Face Liveness." (2014).

[25] Erdogmus, Nesli, and Sébastien Marcel. "Spoofing in 2D face recognition with 3D masks and anti-spoofing with kinect." Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on. IEEE (2013).