https://doi.org/10.2352/ISSN.2470-1173.2017.13.IPAS-208

Face Spoofing Detection Based on Local Binary Descriptors

Yao-Hong Tsai, Yu-Jung Lin; Hsuan Chuang University; Hsinchu City, Taiwan

Abstract

Due to the popularity of the internet, mobile devices have become the main way for delivering messages and online transactions in our daily life. However, facilitate of online transactions will easily derivate security issues. Because of its unique and irreplaceable nature, biometric technology has been widely used in the areas of identification. Among biometrics, face recognition technology for online payments is recently developed. Nowadays, most people still worried about the spoofing of face recognition technology. Under this situations, there are many research results focusing on the important and urgent subjects. Local binary pattern in images is very useful for image processing and computer vision applications like face recognition. Based on the local pattern analysis, we developed a hierarchical analysis algorithm for extracting texture feature such that they not only give a satisfying representation of a face image, but also make the spoofing detection process efficiently. Experimental results show that the proposed method can be applied to a real system for face recognitions.

Introduction

In the rapid development of internet technology, people are frequently logging in their accounts for the web service, such as delivering messages with others and processing online transactions. Person identification technology for various web service is becoming more and more important nowadays. Traditional method for person identification includes ID cards and passwords, but they can be easily divulged to unauthorized users or stolen by impostors. Biometric security systems are becoming preferable since biometrics can provide a natural and convenient ways for person identification by examining the unique and irreplaceable features from the physical or behavioral traits of human beings [1]. Important research results in the literature can be found in the articles [2-3].

With the increasing demand for high-level security in web service through mobile devices, biometric techniques, especially face recognitions, have gained considerable attention recently. Face recognition methods provide a simple and effective method for person identification, requiring only regular cameras and devices. The chairman Jack Ma of Alibaba showed a new face recognition technology for online payments at the computer exhibition in Hanover, German, 2016. It established the relationship between face recognition and online payments, such that face verification method has become the most popular method employed for this task. However, people still worried about the spoofing of face recognition technology. Face spoof attacks usually use a photo or video of an authorized person's face to access the system for facilities or services. Under this situation, there are many research results focusing on the important and urgent subjects, spoofing detection for face recognition [4]. The flow chart of spoofing detection is shown in Figure 1. It can be used to be a preprocess of face recognition systems to determine a captured image is from the real person or not.

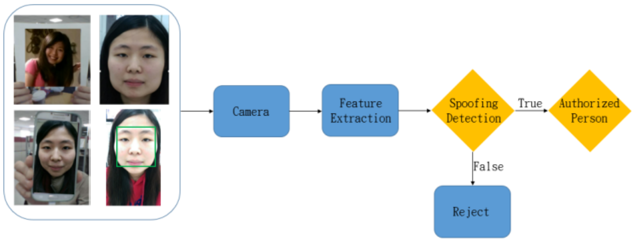

Figure. 1. Spoofing detection for face recognition

While a lot of spoofing detection techniques for face recognition have been proposed recently, their generalization ability has not been adequately addressed since the high-resolution property of handheld mobile devices, such as smart phones and tablets, increased the realism of fake images [5]. Some approaches like motion, spectrum, and image quality are used to solve the problems for spoofing detection in the literature.

For spoofing by printed photos, masks, and screenshots, motion-based approaches are used to develop detection algorithm. Based on the natural responses of the human face, features for spoofing detection including eye blinking [6], [7], mouth movement [8], and head rotation [9] are considered. Pan et al. [6] and Sun et al. [7] detected eye blinking based on the undirected conditional graphical framework. It defined three types of eye states, open, half-open, and close, which is a discriminative measure of eye blinking. Furthermore, Hidden Markov Model and Adaboost algorithm are employed for anti-spoofing. Kollreider et al. [8] proposed face liveness detection based on the optical flow line of the mouth region. They defined the mouth model and extracted the statistics of the lip motion by projected velocity vectors. Finally, they distinguished the real faces and fake ones by Support Vector Machines.

If the video containing the facial gestures, e.g., eye blinking, of the valid user was captured, an attacker could replay it to spoof the security system. For this situation, most motion-based approaches might fail to detect the spoofing. Spectrum-based methods focused on the inter-class difference between live and fake faces by selecting appropriate working spectrums under different wavelength from radiance or reflectance of human face. Kim et al. [9] measured the reflectance disparities between live and fake faces by information from infrared sensor. They computed the radiance under different illuminations and then adopted the Fisher linear discriminant to discriminate these estimated values. Zhang et al. [10] also took advantage of infrared sensor under two different wavelengths to measure the albedo curves of materials from skin and non-skin. Since the above methods all require additional infrared sensor to collect information of faces, it is hard to be a general solution for a standard security system.

Another approach is based on the image quality from face images to detect the fake faces, since the fake ones are usually distorted by the capturing system under the same conditions. Li et al. [11] proposed a method to determine the liveness of the given face image by coefficients of Fourier transform. They assumed that fake faces would lose more details in images by capturing from the face images. Fourier transform are used to extract the high-frequency components from face images. Tan et al. [12] combined texture information and frequency information from the DoG filter to do face liveness detection. Peixoto et al. [13] proposed a standard sparse logistic regression model which is defined as a combination of DoG filters and to overcome the problem under bad illumination conditions. Zhang et al. [14] extracted frequency information by multiple DoG filters and detected fake faces by Support Vector Machines.

The security systems usually rely on flat images in order to recognize human face, so that they can be easily spoofed by photographs or videos of the valid user, which can be easily obtained from the internet or by capturing from the user by using a camera. In this paper, we proposed an efficient and rather robust spoofing detection algorithm for face recognition based on the features of local binary patterns in images.

The contribution of this paper includes the following points:

(1) Overcome the high-resolution property of handheld mobile devices that increased the realism of fake images.

(2) Adapt to environmental changes caused by light illumination instability and simplify computation for real time consideration.

(3) Perform the region based feature extraction by the information of depth and micro texture analysis in images.

(4) Develop efficient face spoofing detection based on extracting the multi-resolution information of the spatial domain in the image plane.

The Proposed Method

In this paper, an efficient and robust spoofing detection algorithm is proposed for face recognition. It extracts the features of hierarchical local binary patterns from a pair of face images. Since local binary pattern in images is very useful for image processing and computer vision applications, many face recognition and related works are developed by the features. Based on the local pattern analysis, we developed a hierarchical analysis algorithm for extracting texture feature including color, texture, and spatial relationship of them. The proposed method is described in the following.

Image acquisition

First, an image acquisition system was designed to capture two images of the user's face with different focal length. The face image capturing system is based on the following assumption. For real human face, the captured images will be partially clear or blur in different focus positions because of depth information of face. On the other hand, they will be all clear or blur for face image display by smart phones or tablets. Sample images of the above two conditions are shown in Figure 2. The left pair of images is captured from the real human face with different focal length such that they result in partially focused situation. On the other hand, the right pair of images is captured from a face image displayed on iPad.

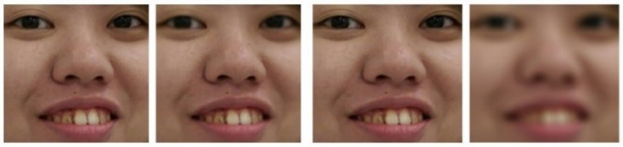

Figure. 2. Sample images of partially focused face image (Left), Sample images of blurred and focused face image (Right)

Features extraction

Then we developed the method for feature extraction by four direction analysis based on image blocks of two images simultaneously. It replaced the original complex direction setting in local pattern analysis. It could also extract the multi-resolution information of the spatial domain in the image plane. Furthermore, we tried to separate the texture and color information by YCbCr color space to generate the local pattern on each image plane and between image planes of Cb and Cr. The proposed method could generate the local pattern according to the spatial information in image plane by the multi-resolution analysis. It could also speed up the computation.

The local binary pattern descriptor [15] labels each pixel of an image with a decimal number, which encode the local structure of the sub-region centered at the pixel. The smallest sub-region in the hierarchy is defined to be a 3x3 image block. For each 3x3 neighborhood, the center pixel is compared with its eight neighbors by subtracting the center pixel value. When the resulting value is negative (positive), it is encoded with 0 (1). The code of the local binary pattern is obtained by concatenating eight resulted 0's or 1's to be a binary number and its corresponding decimal value is used to label the pixel. The above computation is obtained by the following equations. Figure 3 shows an example of local binary pattern.

$$LBP(x,y)=\sum_{p=0}^{P-1} s(f(x,y)-f_p(x,y))2^p, \tag{1}$$

$$s(z)=\begin{cases}1, & \geq 0\\ 0, & <0\end{cases} \tag{2}$$

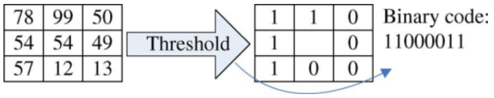

Figure 3. An example of the local binary pattern

Collecting these labels forms the first level of the proposed hierarchical structure. The local binary pattern operates by a sliding window to label pixel by pixel in the image. For the second level, it reduce the local binary pattern by sliding the 3x3 window over two pixels in the resulted labels of first level to generate new codes. According to the similar process, the hierarchical local binary pattern can be obtained level by level and the number of levels depends on the user's requirement. The histogram of hierarchical local binary patterns can be exploited as a texture descriptor for further application of image matching. The final descriptor is constructed by concatenating all the histograms from each level. We then adopt the uniform strategy for pixel weighting to be the approach for descriptor normalization. The diagram of the proposed color spatial local binary pattern (CSLBP) method is shown in Figure 4.

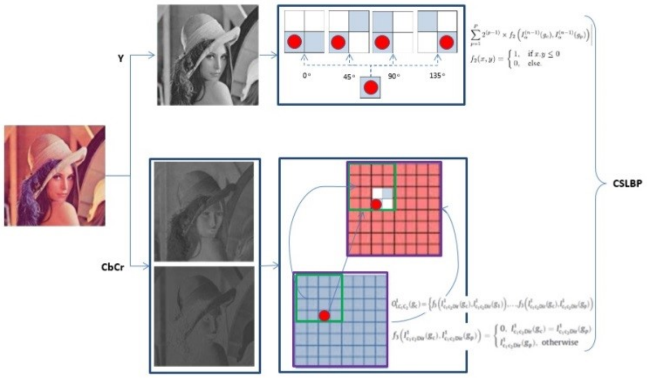

Figure. 4. The Feature generation by CSLBP

The temporary results of the proposed method from illumination correction to features generation are shown in Figure 5. From the first row of this figure, it shows the original image, illumination correction by histogram equalization, gray level image, and the features from CSLBP.

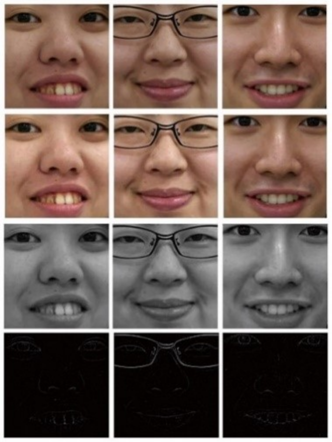

Figure 5. Temporary results of the proposed method

Classifier of Support Vector Machines

Support Vector Machines (SVMs) is a supervised learning models with associated learning algorithms [16]. SVM can be used to analyze data for classification and regression such that it has powerful capabilities in solving pattern recognition for two-class classification problems. Given a set of training examples, an SVM training algorithm builds a model with two categories that can assign new examples to one category or the other. The goal in training a SVM model is to find the separating hyper-plane with maximum distance to the closest points of the training set and minimum training errors. These points are called support vectors. New testing examples are then mapped into that space and predicted to belong to one category or the other based on which side of the hyperplane they fall. The SVM will output the final result of spoofing detection.

Experimental Results

There are several existing benchmarks for spoofing detection like NUAA [19], CASIA-FASD [20], and Idiap REPLAY-ATTACK [21]. Since the proposed method needs to capture two images with different focal length, a new database was built to test the performance. In this experiment, 16 persons in Figure 6 were invited to join the process for online testing and record the video sequences in database. 2000 pairs of images are captured from video sequence and stored in the database. Half of images are used for training to get the threshold for spoofing detection and the other are used to verify the effectiveness of the proposed method.

Figure 6. 16 people for online testing

Face detector of Viola Jones [17] was used to get the face region for online testing system. Figure 7 shows an example of the online testing process. In our experiment, features are input into the linear SVM classifier [18] for training and testing in the last step. The process of online testing was performed for 10 times for each individual subject, which include one test by genuine face and one by iPad to show the face image. Finally, the system was completely succeeded in the experiment.

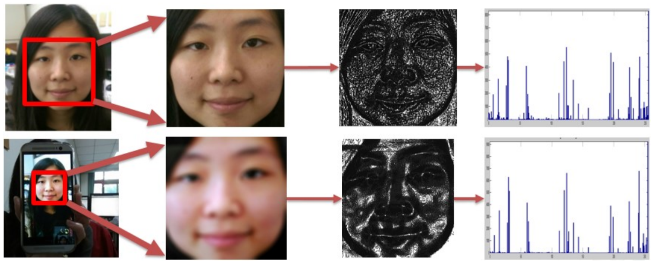

Figure 7. The process for online testing

Experimental results in Figure 8 showed the effectiveness of the proposed descriptors CSLBP in discriminative power and under different illumination changes from 32000 pairs in the database. The normalized value of CSLBP are shown in the left table. It is noted that the threshold values among three sets are quite obvious.

normalization		
Clear	Blur	Partially Focused
0.50	0.12	0.31
0.51	0.13	0.29
0.53	0.13	0.34
0.54	0.14	0.35
0.58	0.12	0.33
0.60	0.13	0.35
0.68	0.15	0.42
0.69	0.14	0.45
0.77	0.16	0.49
0.81	0.16	0.46
0.92	0.18	0.48

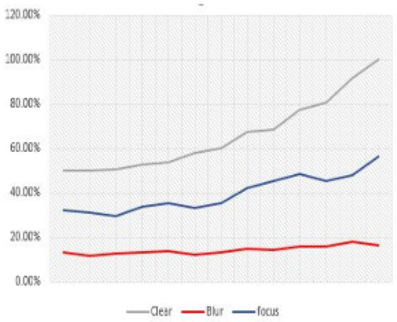

Figure 8. Normalized results of CSLBP features.

Conclusions

Based on the local pattern analysis, we developed a hierarchical analysis algorithm for extracting texture features from two images which were captured with different focal length from the same user. Features of CSLBP not only give a satisfying representation of a face image, but also make the spoofing detection process efficiently. It could overcome the spoof by high resolution images of handheld mobile devices. The method could also adapt to environmental changes caused by lighting illumination instability and simplify computation for real time consideration. Performing the region based feature extraction by the information of depth and micro texture analysis in images resulted in the ability to classify images between authorized users or fake ones. The first part of experimental results showed the efficiency of face spoofing detection based on extracting the multi-resolution information of the spatial domain in the image plane.

References

[1] R. M. Bolle, et al., Guide to Biometrics. Springer Verlag, 2004.

[2] Y. Wang, J. Hu, and D. Phillips, “A Fingerprint Orientation Model Based on 2D Fourier Expansion (FOMFE) and Its Application to Singular-Point Detection and Fingerprint Indexing,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 4, pp. 573–585, 2007.

[3] J. A. Unar, W. C. Seng, and A. Abbasi, “A Review of Biometric Technology Along with Trends and Prospects,” Pattern Recognit., vol. 47, no. 8, pp. 2673–2688, 2014.

[4] G. Pan, L. Sun, Z. Wu, and S. Lao, “Eyeblink-Based Anti-Spoofing in Face Recognition from a Generic Webcamera,” in *Proc. IEEE 11th Int.Conf. Comput. Vis. (ICCV)*, pp. 1–8, 2007.

[5] W. Kim, S. Suh, and J.-J. Han, “Face Liveness Detection From a Single Image via Diffusion Speed Model,” IEEE Trans. on Image Processing, vol. 24, no. 8, 2015.

[6] G. Pan, L. Sun, Z. Wu, and S. Lao, “Eyeblink-based anti-spoofing in face recognition from a generic webcamera,” in *Proc. IEEE 11th Int.Conf. Comput. Vis. (ICCV)*, pp. 1–8, 2007.

[7] L. Sun, G. Pan, Z. Wu, and S. Lao, “Blinking-based live face detection using conditional random fields,” in *Proc. Adv. Biometrics*, pp. 252–260, 2007.

[8] K. Kollreider, H. Fronthaler, M. I. Faraj, and J. Bigun, “Real-time face detection and motion analysis with application in ‘liveness’ assessment,” IEEE Trans. Inf. Forensics Security, vol. 2, no. 3, pp. 548–558, Sep. 2007.

[9] Y. Kim, J. Na, S. Yoon, and J. Yi, “Masked fake face detection using radiance measurements,” J. Opt. Soc. Amer. A, vol. 26, no. 4, pp. 760–766, Apr. 2009.

[10] [13] Z. Zhang, D. Yi, Z. Lei, and S. Z. Li, “Face liveness detection by learning multispectral reflectance distributions,” in Proc. IEEE Int. Conf. Autom. Face Gesture Recognit. (FG), pp. 436–441, 2011.

[11] J. Li, Y. Wang, T. Tan, and A. K. Jain, “Live face detection based on the analysis of Fourier spectra,” Proc. SPIE, Biometric Technol. Human Identificat., pp. 296–303, Aug. 2004

[12] X. Tan, Y. Li, J. Liu, and L. Jiang, “Face liveness detection from a single image with sparse low rank bilinear discriminative model,” in Proc. 11th Eur. Conf. Comput. Vis. (ECCV), pp. 504–517, 2010.

[13] B. Peixoto, C. Michelassi, and A. Rocha, “Face liveness detection under bad illumination conditions,” in Proc. 18th IEEE Int. Conf. Image Process. (ICIP), 2011, pp. 3557–3560, 2011.

[14] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, “A face anti-spoofing database with diverse attacks,” in Proc. IEEE 5th IAPR Int. Conf. Biometrics (ICB), pp. 26–31, 2012.

[15] T. Ojala, M. Pietikäinen and T. Mäenpää, “Multi-Resolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns”, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 24, pp. 971-987, 2002.

[16] V. Vapnik, Slatistical Leoming Theoty, John Wiley and Sons, New York, 1998

[17] P. Viola and M. J. Jones, “Robust real-time face detection,” Int. J. Comput. Vis., vol. 57, no. 2, pp. 137–154, 2004.

[18] C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” ACM Trans. on Intell. Syst. Technol., vol. 2, no. 3, Apr. 2011.

[19] NUAA database http://parnec.nuaa.edu.cn/xtan/data/NUAAImposterDB.html

[20] CASIA-FASD database http://www.cbsr.ia.ac.cn/english/FaceAntiSpoofDatabases.asp

[21] Idiap Replay-Attack database https://www.idiap.ch/scientific-research/resources/replay-attack

Author Biography

Yao-Hong Tsai received the M.S. and Ph.D. degrees in information management from the National Taiwan University of Science and Technology in 1994 and 1998, respectively. He was a Researcher with the Information and Communications Research Laboratories, Industrial Technology Research Institute, Hsinchu, Taiwan. He is currently an Associate Professor with the Department of Information Management, Hsuan Chuang University, Hsinchu, Taiwan. His current research interests include image processing, pattern recognition, cloud computing, and internet of things.