15th IAPR International Conference on Machine Vision Applications (MVA)
Nagoya University, Nagoya, Japan, May 8-12, 2017.

Face Liveness Detection with Feature Discrimination between Sharpness and Blurriness

Chun-Hsiao Yeh, Herng-Hua Chang*
Computational Biomedical Engineering Laboratory
Department of Engineering Science and Ocean Engineering
National Taiwan University, Taipei, Taiwan
danielyehh@gmail.com, herbertchang@ntu.edu.tw

Abstract

Face recognition has been extensively used in a wide variety of security systems for identity authentication for years. However, many security systems are vulnerable to spoofing face attacks (e.g., 2D printed photo, replayed video). Consequently, a number of anti-spoofing approaches have been proposed. In this study, we introduce a new algorithm that addresses the face liveness detection based on the digital focus technique. The proposed algorithm relies on the property of digital focus with various depths of field (DOFs) while shooting. Two features of the blurriness level and the gradient magnitude threshold are computed on the nose and the cheek subimages. The differences of these two features between the nose and the cheek in real face images and spoofing face images are used to facilitate detection. A total of 75 subjects with both real and spoofing face images were used to evaluate the proposed framework. Preliminary experimental results indicated that this new face liveness detection system achieved a high recognition rate of 94.67% and outperformed many state-of-the-art methods. The computation speed of the proposed algorithm was the fastest among the tested methods.

1. Introduction

Over the last decade, information privacy and security has played an important role in human life. Due to this trend, personal authentication has attracted increasing attention and has become a significant issue. Face recognition is one important biometric authentication technique. However, it is particularly vulnerable to attacks that invade the security system and cause damage to personal information. The attacks can be divided into several categories such as printed face photos, recorded videos and 3D mask models with various expressions [1]. In order to address these attacks, researchers have made great efforts to develop anti-spoofing techniques [2].

Depending on different cues used in face liveness detection, existing studies can be categorized into three groups: texture based, motion based and other cues based approaches. In the texture based approach, Määttä et al. [3] applied multi-scale local binary patterns (LBPs) for texture analysis. Li et al. [4] used movement information of live faces based on the analysis of the Fourier spectrum.

In the motion based approach, Sun et al. [5] used eye blinking. Bao et al. [6] applied optical flows to analyze face biometrics. Nevertheless, these studies required long computation time and were easily interfered by background motion.

Other cues can be derived from different sources instead of 2D images. For example, Wang et al. [7] built sparse 3D facial models to identify face liveness. Kim et al. [8] took advantage of the depth information on real faces. They detected real faces using the effect of defocus. Two images respectively focused on the nose and the ears were obtained for feature extraction.

In this paper, we propose a new method to distinguish real human faces from 2D printed face photos. The philosophy underlining the proposed algorithm is that we rely on the property of digital focus with various depths of field (DOFs) while shooting [9, 10]. More precisely, we set the camera focus on the nearest spot of the face, which is the nose. Consequently, for a real face image, the nose will be sharp but other regions such as the cheek will be blurred, depending on the effect of the DOF. On the other hand, the level of clearness for the nose and the cheek will be approximately the same in a 2D printed face image. By analyzing the difference as described above, we can discriminate real face images from printed face images to detect face spoof attacks.

2. Proposed Method

We introduce a new face liveness method based on the digital focus technique to improve existing anti-spoofing systems. Our algorithm is composed of three major procedures, which are face detection and preprocessing, feature extraction, and classification as shown in Fig. 1. The major stages of our face liveness system are summarized as follows:

1. *Face detection and preprocessin*g: In the first stage, we detect the whole face and set the camera focus on the nose. We only analyze the nose located in the center of the face and the cheek located in the bottom right of the face for preprocessing.
2. *Feature extractio*n: We analyze the level of blurriness in both nose and cheek subimages. They should be different due to the effect of the DOF in a real face image. Unlike real human faces, 2D printed face photos are flat, which are expected to have insignificant differences.
3. *Classification*: The k-nearest neighbor (KNN) algorithm for classifying features between real and printed face images is adopted in the final stage. This algorithm is appropriate to classify low dimension feature descriptors as is our case.

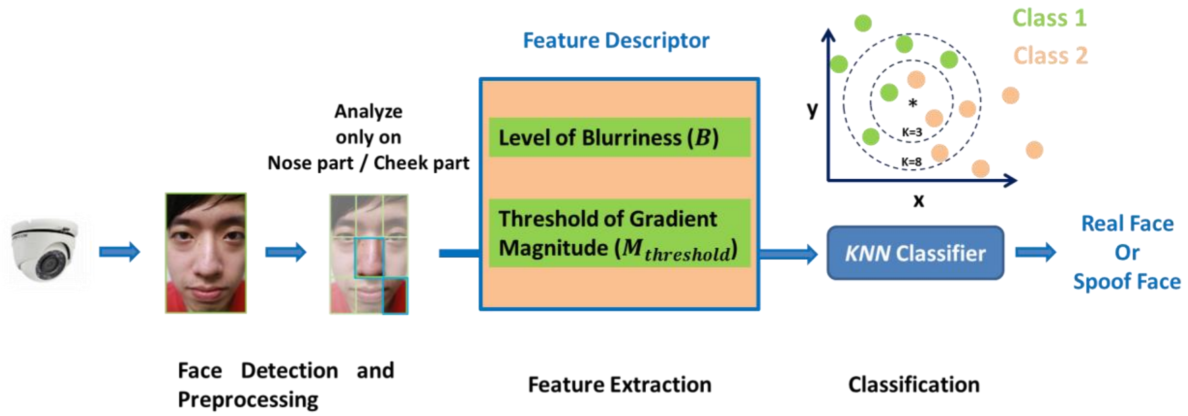
Feature Descriptor
Analyze only on Nose part / Cheek part
Level of Blurriness (B)
Threshold of Gradient Magnitude ($M_{threshold}$)
Class 1
Class 2
y
x
K=3
K=8
KNN Classifier
Real Face Or Spoof Face
Face Detection and Preprocessing
Feature Extraction
Classification

Figure 1. Procedures of the proposed face liveness detection algorithm.

2.1. Face detection and preprocessing

As described in introduction, we detect the whole face as the input image. The camera is set to focus on the nose, which is in the center of the face. The captured nose is supposed to be sharp and clear. However, because of the effect of the DOF, the cheek in the bottom right of a real face image is expected to be blurred. In case of 2D printed face photos, however, there will be no difference between the nose and cheek parts because there is no DOF effect. Therefore, we divide the input face image into 3×3 grids and only extract the nose part and the cheek part. We also downsize both parts by 1/3 in height and width if the size of the input image is too large.

2.2. Feature Extraction

The proposed feature extraction is based on the level of blurriness of the input image. In terms of the gray level, this means that there are strong intensity variations between pixels if the input image is sharp. In contrast, there are weak intensity variations between pixels if the input image is already blurred. The feature extraction procedure is shown in Fig. 2 and described as follows.

First, gradients along both horizontal and vertical directions are computed in the input image $I(x,y)$. The horizontal gradient $F_x(x,y)$ and vertical gradient $F_y(x,y)$ are approximated using the following equations:

$$F_x(x,y) = (-1,0,1) * I(x,y) = \frac{\partial I(x,y)}{\partial x} \tag{1}$$

$$F_y(x,y) = (-1,0,1)^T * I(x,y) = \frac{\partial I(x,y)}{\partial y} \tag{2}$$

where $*$ is the convolution operator. We then use the Gaussian function $G(x,y,\sigma)$ in Eq. (3) to model the blur effect to the input image $I(x,y)$ to obtain $L(x,y,\sigma)$ in Eq. (4).

$$\mathrm{G}(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{3}$$

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \tag{4}$$

Similarly, we compute the horizontal gradient $F_{x\_b}(x,y)$ and the vertical gradient $F_{y\_b}(x,y)$ of the blurred image $L(x,y,\sigma)$ using

$$F_{x\_b}(x,y) = (-1,0,1) * L(x,y,\sigma) = \frac{\partial L(x,y,\sigma)}{\partial x} \tag{5}$$

$$F_{y\_b}(x,y) = (-1,0,1)^T * L(x,y,\sigma) = \frac{\partial L(x,y,\sigma)}{\partial y} \tag{6}$$

The corresponding gradient magnitude images, $M(x,y)$ and $M_b(x,y)$ are accordingly computed using

$$M(x,y) = \sqrt{\left(F_x(x,y)\right)^2 + \left(F_y(x,y)\right)^2} \tag{7}$$

$$M_b(x,y) = \sqrt{\left(F_{x\_b}(x,y)\right)^2 + \left(F_{y\_b}(x,y)\right)^2} \tag{8}$$

To obtain more accurate estimation and reduce the influence of noise, we extract top c percent of pixel points in the gradient magnitude image. More specifically, we set a threshold $M_{threshold}$, which corresponds to the minimum value of top c percent of pixel points using

$$M_{threshold} = \min\left(\text{top } c\% \text{ of } M(x,y)\right) \tag{9}$$

The pixel points that have larger values than the threshold are selected. In other words, we extract higher values of $M(x,y)$ and $M_b(x,y)$ to obtain $S(x,y)$ and $S_b(x,y)$ using

$$S(x,y) = \begin{cases} M(x,y) & \text{if } M(x,y) \geq M_{threshold} \\ 0 & \text{Otherwise} \end{cases} \tag{10}$$

$$S_b(x,y) = \begin{cases} M_b(x,y) & \text{if } M_b(x,y) \geq M_{threshold} \\ 0 & \text{Otherwise} \end{cases} \tag{11}$$

In order to estimate the difference of blurriness between $S(x,y)$ and $S_b(x,y)$, we subtract $S(x,y)$ from $S_b(x,y)$ to generate $D(x,y)$:

$$D(x,y) = S(x,y) - S_b(x,y); \tag{12}$$

where $1 \leq x \leq M$ and $1 \leq y \leq N$ with M and N the dimension of the image. The most significant key is that if the sum of $D(x,y)$ is large, the original image $I(x,y)$ would be sharp. On the other hand, if the sum of $D(x,y)$ is small, the original image $I(x,y)$ would be blurred already.

The difference map $D(x,y)$ is further processed to remove negative values to obtain $V(x,y)$:

$$V(x,y) = \begin{cases} D(x,y) & \text{if } D(x,y) \geq 0 \\ 0 & \text{Otherwise} \end{cases} \tag{13}$$

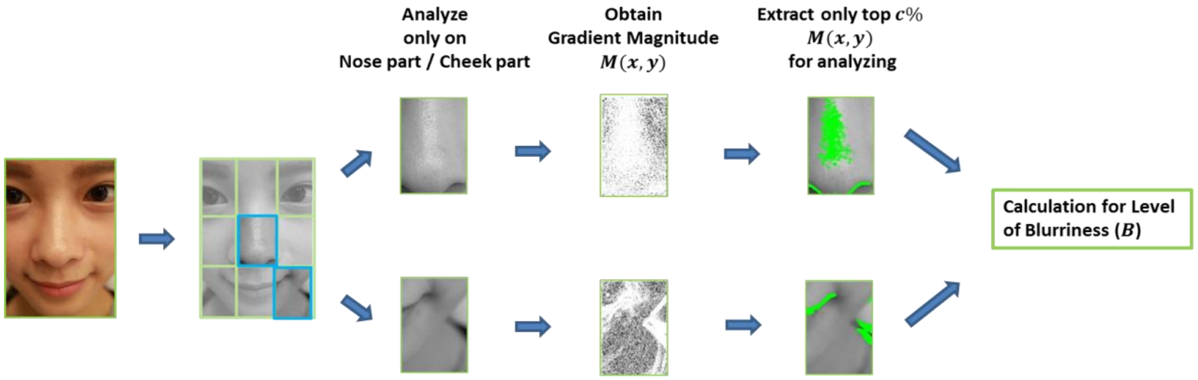
Analyze only on Nose part / Cheek part
Obtain Gradient Magnitude $M(x,y)$
Extract only top $c\%$ $M(x,y)$ for analyzing
Calculation for Level of Blurriness (B)

Figure 2. Procedures of the proposed feature extraction scheme.

We then sum up $M(x,y)$ and $V(x,y)$ pixel by pixel to obtain S_m and S_v as shown in Eqs. (14) and (15), respectively:

$$S_m = \sum_{x=1}^{M}\sum_{y=1}^{N} M(x,y) \tag{14}$$

$$S_v = \sum_{x=1}^{M}\sum_{y=1}^{N} V(x,y) \tag{15}$$

The blurriness level B is defined in terms of S_m, S_v, M, N, and the number of selected pixels using

$$B = \frac{\left(\frac{S_m - S_v}{S_m}\right)\times(MN)}{(\text{number of } c\% \text{ selected pixels})} \tag{16}$$

Finally, the blurriness level B associated with the magnitude threshold $M_{threshold}$ are concatenated together to form the feature descriptor for the following classification process.

2.3. Classification

Finally, the blurriness level B associated with the magnitude threshold $M_{threshold}$ are concatenated together to form the feature descriptor for the following classification process. Specifically, we adopt the KNN algorithm to determine the right class for features extracted from both real and spoofing face images.

3. Experiments

Although there are abundant available databases [11, 12, 13] for liveness detection, the face images are without the specified DOF effects that do not meet our assumptions and requirements. As such, we have established our own database with 75 subjects, each of which had both real human face and 2D printed face images. To systematically evaluate our algorithm, we utilized the following metrics: recognition rate (%) and computation time (second), which measured the accuracy and efficiency of the proposed face liveness system.

3.1. Data Acquisition

Before performing evaluation, we first categorized our database into two groups: training set and testing set. The size of each image is 1920 × 2880 pixels. Each group included real human face and 2D printed photo face images. The images were taken by the Canon EOS M3 with EF-M 22mm lenses. The printed photos for the spoofing dataset were produced by the HP Color LaserJet CP2025 printer. Representative examples of the images are illustrated in Fig. 3.

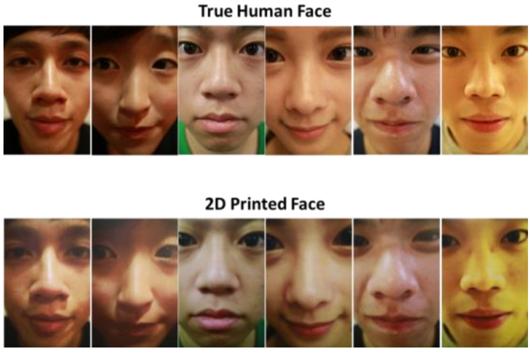

Figure 3. Images selected from the real human face database and the 2D printed photo face database.

3.2. Discrimination Illustration

We then illustrate the discrimination between sharpness and blurriness based on the level of blurriness B and the magnitude threshold $M_{threshold}$. As shown in Fig. 4(a), the differences of B and $M_{threshold}$ between the nose and the cheek were significant in the real face image. This was not the case in the spoofing face image, where the differences of both feature descriptors were insignificant as shown in Fig. 4(b).

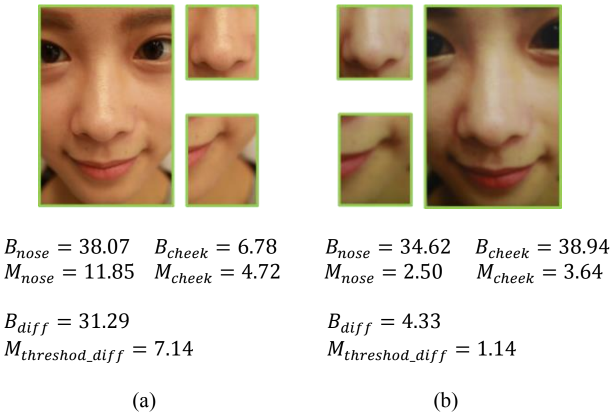
$B_{nose} = 38.07$ $B_{cheek} = 6.78$
$M_{nose} = 11.85$ $M_{cheek} = 4.72$
$B_{diff} = 31.29$
$M_{threshod\_diff} = 7.14$
(a)
$B_{nose} = 34.62$ $B_{cheek} = 38.94$
$M_{nose} = 2.50$ $M_{cheek} = 3.64$
$B_{diff} = 4.33$
$M_{threshod\_diff} = 1.14$
(b)

Figure 4. Illustration of feature discrimination. (a) Real face image. (b) 2D Printed face image.

3.3. Performance Evaluation

The proposed algorithm was compared with three other methods of face liveness detection, including Määttä et al. [3], Wen et al. [14] and Kim et al. [8]. For each input image, we used the nose and cheek parts for feature extraction (B and $M_{threshold}$). A 9 × 9 kernel with σ = 10.5 was used for the Gaussian function throughout the experiments. Besides, $M_{threshold}$ was the minimum value of top 2.14% of pixel points in $M(x, y)$, which was about 1500 pixel points in a 320 × 214 input image.

Seven-fold cross validation experiments were applied to evaluate the performance. As summarized in Table 1, the proposed algorithm achieved TPR=60% @ FAR=0.05 and TPR=75% @ FAR=0.1, which produced average recognition accuracy (%) up to 94.67%. The recognition rate of our algorithm increased 13.12% on average with respect to other existing methods. Table 2 depicts the areas under the receiver operating characteristic (ROC) curves of four methods, which indicated that our proposed scheme accomplished the highest score of 0.9372.

Table 1. Performance comparison of four methods.

Method	**Accuracy (%)**	**TPR @FAR=0.05**	**TPR @FAR=0.1**
Määttä et al. [3]	75.33%	0	0
Wen et al. [14]	82.00%	0.10	0.45
Kim et al. [8]	87.33%	0.60	0.70
Proposed method	**94.67%**	0.60	**0.75**

Table 2. Area under curve (AUC) of four methods.

Method	**Area under curve (AUC)**
Määttä et al. [3]	0.5892
Wen et al. [14]	0.7138
Kim et al. [8]	0.8579
Proposed method	**0.9372**

We also evaluated the computation time required for each method. 120 input images were used for testing. An Intel (R) Core (TM) i5-4570 CPU @ 3.20 GHz machine was used for the experiments. As presented in Table 3, the proposed method reduced on average 50% computation time comparing with existing methods.

Table 3. Runtime (s) of four methods on 120 face images.

Method	Runtime (s)
Määttä et al. [3]	69.732s
Wen et al. [14]	43.464s
Kim et al. [8]	35.755s
Proposed method	**25.596s**

4. Conclusion

In this study, we focused on face liveness detection using digital focus techniques, which efficiently identified facial biometrics. After conducting digital focus with the effect of DOF on images, the properties of real human faces and spoofing faces became apparently different. To efficiently extract features from different parts of the image, we introduced robust feature descriptors for computing the level of blurriness and eventually separated real face images from spoofing face images by the feature classification. We have conducted numerous experiments to evaluate the performance of the proposed algorithm and compared with many state-of-the-art face liveness methods. Preliminary experimental results suggested that the proposed method was effective in identifying face liveness and achieved a 94.67% recognition rate. We believe that there are other features that can be further incorporated into this new face liveness detection framework to make it more secure and robust. More thorough evaluation on parameter settings and large-scale datasets is needed in the future.

Acknowledgments: This work was supported by the Ministry of Science and Technology of Taiwan under Grant No. MOST 105-2221-E-002-082.

References

[1] Schuckers, S.A., Spoofing and anti-spoofing measures. Information Security technical report, 2002. 7(4): p. 56-62.

[2] Galbally, J., S. Marcel, and J. Fierrez, Biometric antispoofing methods: A survey in face recognition. IEEE Access, 2014. 2: p. 1530-1552.

[3] Määttä, J., Hadid, A., & Pietikäinen, M. (2012). Face spoofing detection from single images using texture and local shape analysis. *IET biometrics,* 1(1), 3-10.

[4] Li, J., et al. *Live face detection based on the analysis of fourier spectra*. in *Defense and Security*. 2004. International Society for Optics and Photonics.

[5] Sun, L., et al. *Blinking-based live face detection using conditional random fields*. in *International Conference on Biometrics*. 2007. Springer.

[6] Bao, W., et al. *A liveness detection method for face recognition based on optical flow field*. in *2009 International Conference on Image Analysis and Signal Processing*. 2009. IEEE.

[7] Wang, T., et al. *Face liveness detection using 3d structure recovered from a single camera*. in *2013 International Conference on Biometrics (ICB)*. 2013. IEEE.

[8] Kim, S., et al. *Face liveness detection using variable focusing*. in *2013 International Conference on Biometrics (ICB)*. 2013. IEEE.

[9] Pertuz, S., D. Puig, and M.A. Garcia, *Analysis of focus measure operators for shape-from-focus.* Pattern Recognition, 2013. 46(5): p. 1415-1432.

[10] Pentland, A.P., *A new sense for depth of field.* IEEE transactions on pattern analysis and machine intelligence, 1987(4): p. 523-531.

[11] Tan, X., et al. *Face liveness detection from a single image with sparse low rank bilinear discriminative model*. in *European Conference on Computer Vision*. 2010. Springer.

[12] Zhang, Z., et al. *A face antispoofing database with diverse attacks*. in *2012 5th IAPR International Conference on Biometrics (ICB)*. 2012. IEEE.

[13] Chingovska, I., A. Anjos, and S. 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*. 2012. IEEE.

[14] Wen, D., H. Han, and A.K. Jain, *Face spoof detection with image distortion analysis.* IEEE Transactions on Information Forensics and Security, 2015. 10(4): p. 746-761.