Diffusion speed based face liveliness detection to expose spoofing attack

M.L.Athirai, Dr M.Sujaritha

Abstract— Prevailing face biometric systems are open to spoofing attacks. A spoofing attack occurs when a person tries to masquerade as someone else by misrepresent data and thereby gaining illegitimate access. Spoofing with the help of photographs or videos is one of the most collective approaches of attacking face recognition and highly secured systems. In this a real-time method is proposed based on the diffusion speed the diffusion speed of a single image is obtained to address this problem. Particularly the difference in surface properties between a live face and a fake one (spoofed image) is capably revealed in the diffusion speed, the antispoofing features are exploited by employing the total variation flow scheme. More precisely, defining the local patterns of the diffusion speed is proposed and that is so-called as local speed patterns, as the features, which is the input into the linear SVM classifier to define whether the given face is fake or not. The significant advantage of this method is that, in contrast to previous approaches, it accurately identifies varied malevolent attacks regardless of the medium of the image, e.g., paper or screen. Moreover, this method does not require any specific user cooperation with the system.

Index Terms—spoofing, diffusion speed, total variation flow, local speed pattern, face liveliness detection.

1. INTRODUCTION

With the increasing demand for high-level security in mobile devices, such as smart phones and tablets, biometric techniques have gained substantial attention because of their intrinsic traits. Thus, iris and fingerprint verification systems have been actively researched and are now deployed in various secured systems. But the problem arises where these systems are prone to spoofing attack. Spoofing attack arises where a person masquerades as other person and gains an illegitimate access into a secured system. To address this constraint, there are various methods to discriminate live faces from fake ones such as based on motion, spectrum, and image quality information.

To address the above problems, a novel and simple method for detecting face liveness from a single image is proposed. The key idea of this is that the difference in surface properties between live and fake faces can be efficiently estimated by using diffusion speed. Specifically, computing the diffusion speed by utilizing the total variation (TV) flow scheme and extracting anti-spoofing features based on the local patterns of diffusion speeds, which are so-called as local speed patterns (LSPs). These features are finally given as input into a linear SVM classifier to determine the liveness of the given face image. As compared to previous approaches, this method will perform well regardless of the image medium and even under varying illuminations. This is best suited for achieving robust face recognition and verification system in a wide range of environments under diverse illumination conditions. The rest of this paper is organized as follows. The literature review of existing methods for face liveliness detection in section II. The proposed liveness detection scheme explained in detail in Section III. Implementation of the diffused image with mat lab is explained in section IV & V and the conclusion follows in Section VI.

2. LITERATURE REVIEW

This section explains the existing methods for face liveliness detection to eradicate spoofing attacks.

1.1. Local Binary Pattern based anti-spoofing

In this method the LBP histogram is calculated in two different ways, and all the experiments are performed separately on the both versions of the feature vectors. The first option is to calculate the LBP features for all the pixels in the image and distribute them in one histogram (per-image calculated features). In this case, the total number of bins in the histogram, and the number of dimensions of the feature vector is 59. The second option is to divide the image into 3x3 blocks, calculate the LBP histograms for each of the blocks separately and form the final feature vector by their concatenation (per-block computed features). This results in a feature vector with 531 dimensions. The advantage of using blocks comes from the fact that the texture artifacts of the spoof attacks may be more visible in small and local uniform areas of the image, such as the forehead or the cheeks [1]. Then the feature vectors of the probe images are assigned a score and these features are given to a classifier. More complex classifiers are examined as well a linear one, Linear Discriminant Analysis (LDA) and a non-linear one, Support Vector Machine (SVM) with radial kernel basis function. Finally the LBP based anti-spoofing method guarantees different levels of certainty for different types of attacks and different databases. Some attacks can step out this counterfeit more easily than others. There is no consistency in the results with regards to the types of attacks, nor the attacks from different databases.

1.2. OFC based methods for photo attacks

In this a new technique of countermeasure is proposed merely based on foreground/background motion correlation using Optical Flow and achieving nearly perfect scoring with

an equal-error rate of 1.52% on the available test data. This algorithm tries to detect motion correlations between the head of the user trying to authenticate and the background of the scene, which indicates the presence of a spoofing attack [2]. Instead of working with averaged intensities, it uses fine-grained motion direction for deriving the correlation between foreground and background region.

The direction of objects in the scene is estimated using Optical Flow (OF) techniques. The use of OF is expected to grant more precise estimation of motion parameters between the regions of interest in the scene, assuring that motion signs are related in direction and do not come from unrelated phenomena. Instead of lump-summing intensities, OFC (Optical flow correlation) quantizes, histograms, normalizes and directly compares motion direction vectors from the two regions of interest in order to provide a correlation score, for every analyzed frame.

OFC also introduces a new hyper-parameter that controls the amount of specific or global information that is considered while performing discrimination. As, the number of directions Q used by the algorithm determines if the detector will observe motion patterns which may be related to specific acquisition conditions or application independent. This method has disadvantage that it is vulnerable to video attacks.

1.3. Lambertian model

Usually, natural media in which fake faces exist primarily include paper, screen of video device, and photo and so on. The structure of these media is greatly different from that of live face. All these media are 2-D planar structure, whereas live face is 3-D structure. According to the Lambertian model [3], the face image can be described as,

$$I(x,y)=\rho(x,y)n(x,y)Ts \quad (1)$$

Where ρ is the albedo (surface texture) of face, n(x, y)T is the surface normal (3D shape) of the object (the same for all objects of the class), and s is the point source, which can vary arbitrarily. Due to 2-D planar structure of photograph, n(x,y)T is a constant. So, under the same illumination, images from a liveness are determined by the albedo and the surface normal, whereas those from a fake are determined by the albedo only. Thus a conclusion can be draw from equation (1) that the intensity contrast of live face image is more obvious than that of fake image. Such differences lead to their greatly different reflectivity of light, which is reflected in frequency distribution of an image. Additionally, the size of fake image is usually smaller than that of live face. If they are held before the camera, many details contained in the face captured by the Camera will lose. All these bring about great differences between live face image and fake face image, which can be detected by analyzing their 2D Fourier spectra. This method is constrained only to a small set of database. Table 1 is the comparison of existing algorithms HTER is the Half the Total Error Rate.

Table I. Comparison table for the existing system

Methods	LBP	OFC	Lambertian
HTER	8.98%	89%	7.9%
Time	33.2 msec	33.4 msec	30.7

3. PROPOSED METHOD-LIVELINESS DETECTION

The various stages of the proposed method of face liveliness detection is given in Figure1. The input image given may be an original image or spoofed image and the image undergoes five stages of processing. Initially preprocessing include finding of diffused image from the input image.

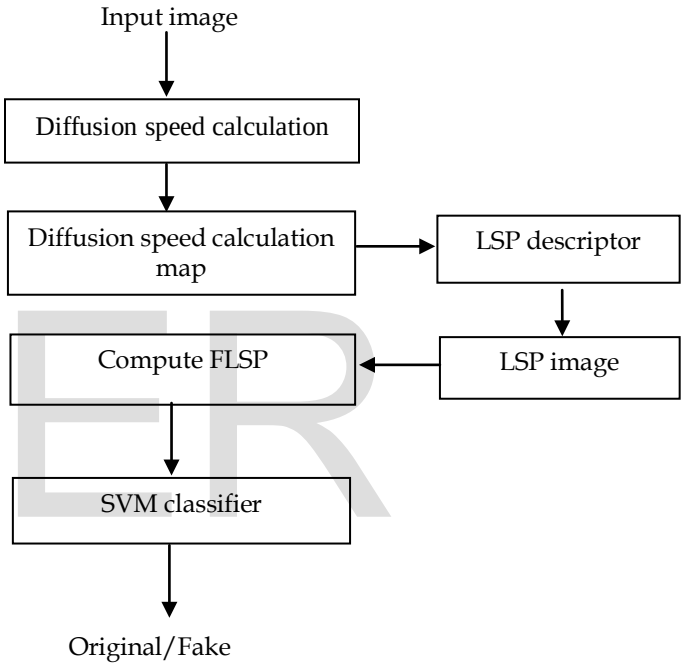

Figure1: Proposed System architecture

Therefore, it is considered that the diffusion speed, e.g., the difference in pixel values between the original and diffused images, provides useful clues that can be used to discriminate a live faces from a fake one in a single image. In particular, it is attempted to model this diffusion process by allowing for the total variation (TV) flow scheme, and extract anti-spoofing features based on the local patterns of the diffusion speed values computed at each pixel position. In the following, the proposed method is explained in detail. The HTER values of the proposed method for development and test on the whole set are 13.72% and 12.50%, respectively.

3.1. Diffusion Speed

In this subsection, the diffusion Speed is shown in which illumination characteristics are clearly revealed. To this end, nonlinear diffusion on the original face image I is conducted given as [4],

$$u_{k+1}=u_k+\operatorname{div}\left(d(|\nabla u_k|\nabla u_k\right),\ u(k=0)=I \quad (2)$$

Where k denotes the iteration number. For the diffusivity function d(.), adopting the total variation (TV) flow is, defined as [5],

$$d(x) = 1/(x+\xi) \quad (3)$$

Where ξ is a small positive constant. An important issue is to solve the diffusion equation defined in (1). To this end, efficient approach is used, called the additive operator splitting (AOS) scheme [6] defined as,

$$u_{k+1}=\tfrac{1}{2}\left((I-2\tau A_x(u_k))^{-1}+(I-2\tau A_y(u_k))^{-1}\right) \quad (4)$$

where Ax and Ay denote the diffusion matrices computed in the horizontal and vertical directions, respectively this AOS scheme is unconditionally stable, and thus, it is possible to use a large time step, e.g., τ = 30, which provides a good compromise between efficiency and accuracy, to enable fast diffusion.

3.2. Diffusion Speed Map

On the basis of the above analysis, the ability of the diffusion speed model to efficiently extract anti-spoofing Features is utilized. More specifically, straightforwardly employ the value of the diffusion speed itself at each pixel position as the baseline features, given as

$$F_{base} = \{s(x, y) \mid 0 < x \le W, 0 < y \le H\} \quad (5)$$

Where W and H denote the width and height of the detected face region, respectively. The local speed patterns to efficiently capture even small differences between the diffusion speed maps of live and fake faces. Thus, the range of fLSP(x, y) is [0, 255] and can be represented as a gray-scale image (LSP image).

3.3. LSP Descriptor

The original face is given and the Diffusion speed map is scaled from [0, 255]. In the face image obtained the dark color indicates faster moving pixels and then fLSP at each pixel position is computed finally an LSP image is obtained. To ensure a fair comparison, it is been applied the LSP operator to the results of the LTV diffusion to generate the feature vector, which was input into the linear SVM classifier. As can be seen, this LSP-based features achieve the best performance, 98.5%, using only the simple linear classifier.

3.4. Compute FLSP

LSP-based feature vector generation for the given face image is generated the dimension of the proposed feature vector is 59 × K, where K is the number of image blocks and FLSP is obtained.

3.5. LSP Image

In this method diffusion speed is focused rather than the diffusion result itself, as in the logarithmic total Based on the TV flow-based diffusion speed, which is quite different from the traditional total variation framework used in the LTV model, this method can efficiently reveal the difference in the reflectance characteristics according to the 2D plane and 3D structure, whereas the LTV model provides only the illumination-invariant face image, regardless of the liveness of the given face . As compared to the texture patterns widely employed in previous approaches, this LSP-based feature vector captures illumination characteristics on corresponding surfaces. This allows the proposed scheme to be robust to a wide range of spoofing attacks using various media. Moreover, it has a very good ability to discriminate live faces from fake ones, even when the latter are captured in high resolution.

3.6. Support Vector Machine (SVM) Classifier

The idea behind SVMs is to make use of a (nonlinear) mapping function θ that transforms data in input space to data in feature space in such a way as to render a problem linearly separable. The SVM then automatically discovers the optimal separating hyperplane (which, when mapped back into input space via θ-1, can be a complex decision surface). SVMs are rather interesting in that they enjoy both a sound theoretical basis as well as state-of-the-art success in real-world applications. In this section the difference in surface properties between live and fake faces which is obtained using diffusion speed and Specifically, computing the diffusion speed by utilizing the total variation (TV) flow scheme and extracting anti-spoofing features based on the local patterns of diffusion speeds, the so-called local speed patterns (LSPs). These features are subsequently input into a linear SVM classifier to determine the liveness of the given face image.

4. DATASETS

This dataset [10], which is the most widely adopted benchmark for the evaluation of face liveness detection, comprises images of 15 subjects who were asked to frontally look at the webcam (capturing faces at 20 fps) with a neutral expression. In addition, none of the faces contains any apparent movement, such as eye blink or head movement. To create fake examples, the authors of [10] captured pictures of each subject using a usual Cannon camera and printed them on photographic paper and normal A4 paper, respectively.

All the faces were detected by using a Viola-Jones detector [11] and geometrically normalized based on the eye localizer [12]. Finally, these images were resized to 64 x 64 pixels with gray-scale representation. Some samples of the NUAA dataset are shown in Fig. 2. For the training set, a total of 3,491 images (live: 1,743 / fake: 1,748) were selected, while the test set was composed of 9,123 images (live: 3,362 / fake: 5,761). It should be noted that there is no overlapping between the training and test sets.

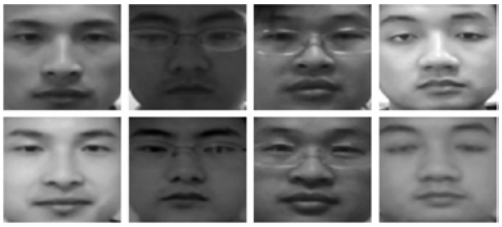

Figure 2 Samples from the NUAA dataset (top: live faces; bottom: fake faces).

5. EXPERIMENTAL RESULT

The simulation scenario is created where the input image is given from NUAA dataset [10]. Following figures is the output obtained. SVM yield high performance on the development set, but are less effective on the test set. This can be explained by the fact that the classification threshold is chosen on the development set, which for NUAA is actually a subset of the training set, as we perform cross-validation. This problem can be taken as an indication for the necessity of a precise protocol with separate training, development and test set in spoof-attack databases.

(a)Input image

(b) Diffused image

(c) Diffusion map

Table II . HTER for LSP based face detection

Method	LSP
HTER	13.72
Time	33 msec

6. CONCLUSION AND FUTURE WORK

A simple and robust method for face liveness detection is proposed in this paper. The key idea of the proposed method is to adopt diffusion speed for demonstrating the difference in the illumination characteristics of live and fake faces. Specifically the TV flow and AOS scheme is exploited to capably compute the diffusion speed, which is robust to changing lighting conditions. To capture the difference between live and fake faces more effectively, an endeavor is made to encode the local pattern of diffusion speed values, the so-called local speed pattern (LSP) which is input to SVM classifier. Thus the proposed method successfully performs when the images are captured in a wide range of indoor and outdoor environments, and when it include persons with varying poses and expressions and under different illuminations. To improve the processing of images under unsupervised learning neural networks can be used as classifier. Moreover, this LSP-based scheme is effective in real-time and can thus be deployed in various mobile devices. Therefore the proposed method for face liveness detection will lead to high-level security for mobile devices.

REFERENCES

[1] J. M¨a¨att¨a, A. Hadid, and M. Pietik¨ainen, "Face spoofing detection from single images using micro-texture analysis," in Proc. International Joint Conference on Biometrics (IJCB 2011), Washington, D.C., USA, 2011.

[2] Anjos and S. Marcel, "Counter-measures to photo attacks in face recognition a public database and a baseline," in International Joint Conference on Biometrics (IJCB), 2011.

[3] B.Clarkson, T.Jebara, A.Pentland, "Multimodal PersonRecognition using Unconstrained Audio and Video", In Proceedings of the 2nd International Conference on Audio-Visual Biometric Person Authentication, 1998.

[4] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, no. 7, pp. 629–639, Jul. 1990.

[5] M. Rousson, T. Brox, and R. Deriche, "Active unsupervised texture segmentation on a diffusion based feature space," in Proc. IEEE Comput. Soc. Comput. Vis. Pattern Recognit. (CVPR), vol. 2. Jun. 2003, pp. II-699–II-704.

[6] J. Weickert, B. M. T. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," IEEE Trans. Image Process., vol. 7, no. 3, pp. 398–410, Mar. 1998.

[7] T. Chen, W. Yin, X. S. Zhou, D. Comaniciu, and T. S. ation models for variable lighting face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 9, pp. 1519–1524, Sep. 2006

[8] Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in Proc. IEEE Int. Conf. Biometrics Special Interest Group (BioSIG), Darmstadt, Sep. 2012, pp.1–7.

[9] 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), 2010, pp. 504–517.

[10] P. Viola and M. J. Jones, "Robust real-time face detection," Int. . 57, no. 2, pp. 137–154, 2004.

[11] X. Tan, F. Song, Z.-H. Zhou, and S. Chen, "Enhanced pictorial structures for precise eye localization under uncontrolled conditions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2009, pp. 1621–1628.