PPGSecure: Biometric Presentation Attack Detection Using Photopletysmograms

Ewa Magdalena Nowara, Ashutosh Sabharwal, Ashok Veeraraghavan
Rice University
6100 Main St, Houston, TX 77005
[emn3, ashu, vashok] @rice.edu

I. Introduction

Using face recognition as a form of authentication has become widespread due to the advances in face detection and recognition algorithms. The existing state of the art face recognition algorithms can recognize a given user with 99% accuracy [1]. Biometrics authentication systems based on face recognition are already commonly used in applications ranging from border security to unlocking smartphones. Despite being commonly used and their high recognition accuracy, face recognition algorithms suffer from vulnerability to simple spoofing attacks. For instance, an attacker may easily obtain a photograph of the authentic user by downloading it from their social media page and use it to successfully fool the face recognition system. In addition to photographs, more sophisticated methods of attacks, such as replaying a video of the user or making a realistic 3D mask have been used [2].

In this work we propose using physiology as a new method of verifying liveness of a user's face. Using a regular camera we can observe photopletysmograms (PPG) which are signals related to color changes in the skin caused by blood flow. We can differentiate between a live face and a face attack or the background by training a machine learning classifier on the frequency spectra of these PPG signals. Machine learning is able to pick out subtle patterns present in the frequency spectra of live signals and accurately classify a presented face as live or as an attack.

In addition, we consider how this work can be extended to detecting live skin regions other than the face. For example, this can aid in finding survivors during a rescue mission by using drones with cameras. Finally, we consider challenges and limitations of the proposed method and possible improvements which remain to be done in the future.

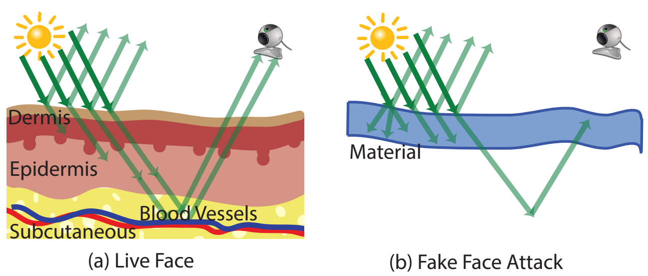

Fig. 1. PPG signals derived from color changes due to blood flow can be observed from a video recording of a live face because some of the light is able to pass through the skin and reach blood vessels. These types of color changes are not present in face attacks because there are no blood vessels present. Therefore, the observed intensity changes do not have the characteristic PPG signals properties.

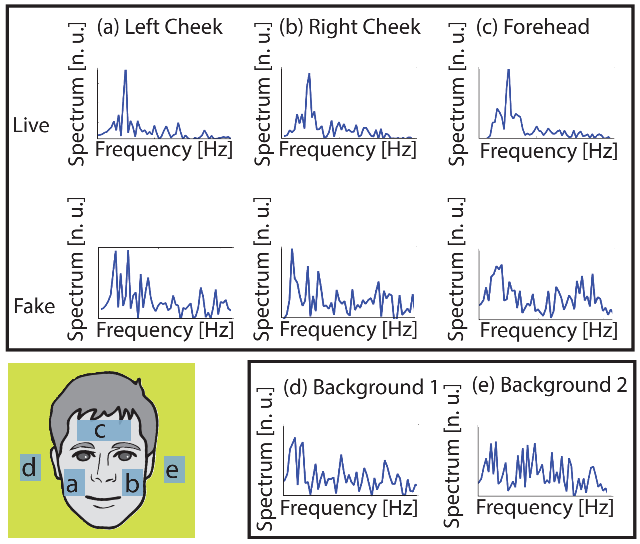

Fig. 2. PPG signals from different facial regions on a live face share characteristic similarities that are abscent in signals from a face attack or the background regions.

II. Overview

An RGB camera can detect PPG signals caused by blood flowing through the circulatory system of a live skin region. These PPG signals are absent in areas which do not contain live skin regions. Some light passes through the skin and some is reflected at the surface. A portion of the light that passes through the skin is absorbed at the surface in the dermis skin layer by melanin present in the epidermis layer, and some remaining light reaches the blood vessels. When a material covers the skin, the majority of the light is absorbed or reflected by that material and only a small portion of the light reaches the skin beneath. Therefore, the captured video does not contain these subtle pulsatile color changes induced by the blow flow (See Figure 1). Furthermore, signals from several facial regions share similarities in the frequency spectra and have a peak related to a heart beat frequency, around

1 Hz band. Signals measured from the background and from face attack materials, such as photographs or videos, have random frequency spectra without these common similarities. This allows us to detect a difference between a live skin region and other elements in the scene. We illustrate the differences in the observed frequency spectra of PPG signals from live faces and face attacks in Figure 2.

III. Prior Work

Various anti-spoofing methods have been explored to prevent attacks on face recognition systems using a fake face, such as a photograph, video or a mask [2]. Prior anti-spoofing techniques can be categorized as motion-based, appearance-based [2] and physiology-based.

A. Motion and Appearance Based Anti-spoofing

Motion-based techniques consider the differences in motion between live authentic faces and face attacks, such as blinking [3]–[5], gaze [6] or pupillary reflex [7], [8]. Meanwhile, some appearance-based methods used differences in texture and spectral reflectance between live faces and face presentation attacks [9]–[11], as well as differences in multispectral properties of skin and mask materials [12]. These methods are designed to prevent a specific kind of attacks and they often don't generalize well to many different kinds of sophisticated attacks. For example, if an attacker is wearing a mask made of a realistic material with holes cut out for their eyes, neither the motion nor texture based methods will be able to classify it as an attack.

B. Physiology Based Anti-spoofing and Liveness Detection

A recent approach employed by several groups is to use camera-based physiology measurements to design an anti-spoofing technique. Due to rapid advances in camera-based vital signs detection, such as pulse rate, pulse rate variation and breathing rate [13]–[17], it is possible to use a regular webcam to detect PPG signals related to blood flow in the skin. Since those PPG signals detected from live skin regions share properties that differentiate them from other signals, several approaches used this property to for liveness detection. The goal of liveness detection is to locate the live skin regions in the videos, while the goal of anti-spoofing methods is to verify that a presented face corresponds to a live authentic user. Existing attempts in the literature of physiology-based anti-spoofing or liveness detection are limited to datasets with a small variety of attacks or do not address the more challenging issues of varying light conditions and motion [18]–[21].

IV. Methodology of PPGSecure

The algorithm we developed to distinguish between live faces and face attacks is called PPGSecure. First, we extract PPG signals from the forehead and the cheeks, as well as from the background regions behind the person. We select these particular facial regions becasue the PPG signals tend to be the strongest in those areas. The advantage of including the background regions in the spectral features is that any temporal variations in intensity induced due to illumination intensity fluctuations will be the same for the face in the foreground and the background regions. But the physiological pulsatile signals will induce intensity changes only in a live face in the foreground.

To compute the PPG signals, we detect facial landmarks [22], track the facial regions of interest [23] and average the temporal intensity changes in the green channel to obtain a single PPG signal describing each region of interest. Once we have extracted the raw PPG signals from the face and the background, we subtract the mean and bandpass filter the PPG signals in physiological range, [0.5 Hz, 5 Hz]. The magnitude of the Fourier spectrum of each filtered PPG signal is a spectral feature. We concatenate these spectral features from three facial regions and two background regions to obtain a spectral feature vector for classification.

We train a support vector machine (SVM) [24] and a random decision forest (RDF) classifier [25] on these spectral features of training subjects' videos. We use a leave-one-subject-out validation method to avoid training and testing on spectral features from videos of the same person.

V. Results

PPGSecure outperforms the state-of-the-art [19] on a publicly available Replay-Attack dataset which contains photograph and video attacks, both fixed and handheld in front of the camera. Liu et al. 's [19] performance drops when the face attack is handheld in front of the camera. Their performance was 88% on photograph attacks and 85% on video attacks, compared to 99% - 100% accuracy of PPGSecure. This could be because Liu et al. 's method looks for correlated changes in the facial regions and handshake motion makes the whole photograph or video move uniformly, resulting in strong cross-correlation patterns. This work, PPGSecure performs better when the PPG signals are bandpass filtered before taking the Fourier transform, improving the initial result from 83% accuracy on video attacks and 91% on photograph attacks to 99% accuracy. This could be because bandpass filtering removes unrelated noise from frequency bands outside the physiological range. Furthermore, adding background regions improves the performance of PPGSecure resulting in close to 100% final accuracy.

VI. Limitations and Challenges

The accuracy of the proposed physiology-based liveness detection method is limited by many factors influencing the accuracy of physiology measurements from a video recording. Some of these limitations are listed below.

A. Skin partially covered

In this work we assumed that a face is fully visible in the camera. However, there are cases when a person has facial hair covering the skin regions. For this method to be extended beyond face anti-spoofing applications, we need to consider a situation when there are several regions of the skin visible and several covered regions. We will need to develop a strategy to consolidate these signals from different regions to determine where a person is located.

B. Motion and varying lighting conditions

When a person is moving, it leads to sudden changes in the incident illumination on the skin and shadows, corrupting the PPG signal measurement. Furthermore, varying illumination will cause difficulties, for example when a person is outdoors. PPG measurements are based on temporal, periodic intensity variations in the skin, therefore any unanticipated intensity changes not related to the PPG signal will make it difficult to detect the PPG signal.

C. Distance

PPG signals extracted from a video recording are very weak, therefore, when a camera is located far from the skin region, PPG signals are harder to detect.

VII. CURRENT AND FUTURE WORK

The described PPGSecure method did not adress the mentioned difficulties but understanding and overcoming these challenges is a part of our future work. We are currently exploring the aforementioned challenges and developing a more accurate method to detect and amplify the weak PPG signal. At the same time our method should not remove all of the noise so that is possible to distinguish between living skin and the background. Future work includes improvements in PPG signal detection, as well as extending this method beyond face anti-spoofing to live skin detection.

A. Robust PPG measurement

Simple averaging of the intensities to get a rough PPG signal, as it was done in PPGSecure may not be sufficient in more difficult scenarios. Obtaining a more accurate PPG signals measurement can aid in improving liveness detection in low light, motion and larger distance. We are working on understanding the theoretical maximum distance and motion and minimum light conditions for the proposed method to work.

B. Extension To Live Skin Detection

Liveness detection does not require an exact measurement of the vital signs but a rough measurement of the PPG waveform may be sufficient to distinguish live regions from fake ones. We are working on understanding if this method can work in more adverse scenarios than methods intended for health monitoring. We are working on analyzing what is the maximum distance that a person can be from the camera for this method to work. Additionally, we are exploring the minimum camera requirements, such as temporal and spatial resolution, as well as how much motion can be present between the person and the camera.

REFERENCES

[1] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 815–823.

[2] J. Galbally, S. Marcel, and J. Fierrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530–1552, 2014.

[3] H.-K. Jee, S.-U. Jung, and J.-H. Yoo, "Liveness detection for embedded face recognition system," *International Journal of Biological and Medical Sciences*, vol. 1, no. 4, pp. 235–238, 2006.

[4] G. Pan, L. Sun, Z. Wu, and S. Lao, "Eyeblink-based anti-spoofing in face recognition from a generic webcamera," in *2007 IEEE 11th International Conference on Computer Vision*. IEEE, 2007, pp. 1–8.

[5] J. W. Li, "Eye blink detection based on multiple gabor response waves," *Proceedings of the 7th International Conference on Machine Learning and Cybernetics, ICMLC*, vol. 5, no. July, pp. 2852–2856, 2008.

[6] A. Ali, F. Deravi, and S. Hoque, "Liveness detection using gaze collinearity," *Proceedings - 3rd International Conference on Emerging Security Technologies, EST 2012*, pp. 62–65, 2012.

[7] X. Huang, C. Ti, Q. Z. Hou, A. Tokuta, and R. Yang, "An experimental study of pupil constriction for liveness detection," *Proceedings of IEEE Workshop on Applications of Computer Vision*, pp. 252–258, 2013.

[8] A. Pacut and A. Czajka, "Aliveness detection for iris biometrics," in *Proceedings 40th Annual 2006 International Carnahan Conference on Security Technology*. IEEE, 2006, pp. 122–129.

[9] I. Chingovska, 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*. IEEE, 2012, pp. 1–7.

[10] T. Pereira F., J. Komulainen, A. Anjos, J. Martino M., A. Hadid, M. Pietikainen, and S. Marcel, "Face liveness detection using dynamic texture," *EURASIP Journal on Image and Video Processing*, p. 2, 2014.

[11] J. Määttä, A. Hadid, and M. Pietikäinen, "Face spoofing detection from single images using micro-texture analysis," in *Biometrics (IJCB), 2011 international joint conference on*. IEEE, 2011, pp. 1–7.

[12] N. Kose and J. L. Dugelay, "Reflectance analysis based countermeasure technique to detect face mask attacks," *2013 18th International Conference on Digital Signal Processing, DSP 2013*, vol. 1, pp. 0–5, 2013.

[13] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 7–11, 2011.

[14] Y. Sun and N. Thakor, "Photoplethysmography revisited: from contact to noncontact, from point to imaging," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 463–477, 2016.

[15] D. McDuff, S. Gontarek, and R. W. Picard, "Improvements in remote cardiopulmonary measurement using a five band digital camera," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 10, pp. 2593–2601, 2014.

[16] M. Kumar, A. Veeraraghavan, and A. Sabharwal, "Distanceppg: Robust non-contact vital signs monitoring using a camera," *Biomedical optics express*, vol. 6, no. 5, pp. 1565–1588, 2015.

[17] S. Tulyakov, X. Alameda-Pineda, E. Ricci, L. Yin, J. F. Cohn, and N. Sebe, "Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2396–2404.

[18] K. H. Suh and E. C. Lee, "Face liveness detection for face recognition based on cardiac features of skin color image," in *First International Workshop on Pattern Recognition*. International Society for Optics and Photonics, 2016, pp. 100 110C–100 110C.

[19] S. Liu, P. C. Yuen, S. Zhang, and G. Zhao, "3d mask face anti-spoofing with remote photoplethysmography," in *European Conference on Computer Vision*. Springer, 2016, pp. 85–100.

[20] S. Bobbia, Y. Benezeth, and J. Dubois, "Remote Photoplethysmography Based on Implicit Living Skin Tissue Segmentation," 2016.

[21] W. Wang, S. Stuijk, and G. De Haan, "Unsupervised subject detection via remote PPG," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2629–2637, 2015.

[22] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1867–1874.

[23] B. D. Lucas, T. Kanade *et al.*, "An iterative image registration technique with an application to stereo vision." in *IJCAI*, vol. 81, no. 1, 1981, pp. 674–679.

[24] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[25] T. K. Ho, "Random decision forests," in *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, vol. 1. IEEE, 1995, pp. 278–282.