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

***Abstract*— Authentication of users by exploiting face as a biometric is gaining widespread traction due to recent advances in face detection and recognition algorithms. While face recognition has made rapid advances in its performance, such face-based authentication systems remain vulnerable to biometric presentation attacks. Biometric presentation attacks are varied and the most common attacks include the presentation of a video or photograph on a display device, the presentation of a printed photograph or the presentation of a face mask resembling the user to be authenticated. In this paper, we present PPGSecure, a novel methodology that relies on camera-based physiology measurements to detect and thwart such biometric presentation attacks. PPGSecure uses a photoplethysmogram (PPG), which is an estimate of vital signs from the small color changes in the video observed due to minor pulsatile variations in the volume of blood flowing to the face. We demonstrate that the temporal frequency spectra of the estimated PPG signal for real live individuals are distinctly different than those of presentation attacks and exploit these differences to detect presentation attacks. We demonstrate that PPGSecure achieves significantly better performance than existing state of the art presentation attack detection methods.**

I. Introduction

Authentication systems using biometrics are already commonly used in a variety of applications, ranging from mobile phones to border security, because they are easy to use and provide a potentially higher level of security. Instead of memorizing a lengthy password that could be intercepted by a hacker, the user only needs to use their finger or their face to confirm their identity. Despite being commonly used, these biometrics-based authentication systems are still vulnerable to spoofing attacks where an attacker can gain access to the user's unique biometric.

A biometric presentation attack (BPA) is a situation in which an attacker has obtained the authentic user's biometric and is using it to fool the biometrics-based authentication system to access the user's devices and accounts. For example, by downloading a picture or a video of the user from their social media page, the attacker may be able to fool the system that relies on face recognition. There have even been cases where attackers 3D print facial masks or fingerprints and can successfully *spoof* the authentication system [1].

We developed PPGSecure, a physiology-based biometric presentation attack detection (BPAD) algorithm which determines whether a face presented to a biometrics authentication system is alive or if it is a face BPA, such as a photograph or a video of the user. Figure 1 shows an overview of PPGSecure liveness detection. PPGSecure detects a photoplethysmogram (PPG), which is a signal caused by small color changes in the skin due to the blood flow. These PPG signals contain physiological indicators that are observable only in videos of alive faces, allowing machine learning models to accurately classify a presented face as live or an attack.

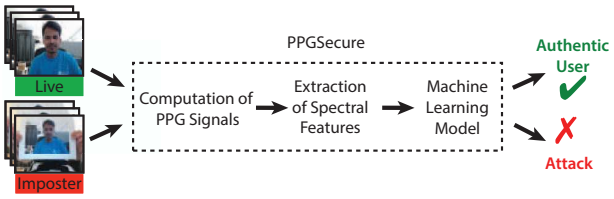

Fig. 1. Overview of PPGSecure. Frequency analysis of PPG signals extracted from captured images is used to distinguish live users from spoofing attacks with photographs or videos.

The novelty of our approach is that we rely on generic frequency features of the entire frequency spectra filtered in the physiological frequency range, instead of choosing specific frequency bins or properties of the spectra, such as the location of the maximum peak [17], [30]. We use machine learning algorithms to find discriminative patterns in the frequency spectra that may be difficult to notice by a human. The advantage of using the entire frequency spectra directly makes PPGSecure robust to a variety of attacks because we do not have to design what signal features might be discriminative of real live faces which may vary for different methods of fraud.

The paper is organized as follows. Related work on face anti-spoofing and camera-based PPG detection is described in Section 2. In Section 3, we provide an overview of the proposed idea to give intuition about why it works, followed by the details of our proposed algorithm in Section 4. We report our results in Section 5 and offer comments about possible improvements and sources of error in Section 6.

II. Prior Work

A. Face Presentation Attacks

Spoofing attacks used to fool facial recognition systems have been identified as either 2-dimensional - printed pho-

978-1-5090-4023-0/17 $31.00 © 2017 IEEE
DOI 10.1109/FG.2017.16

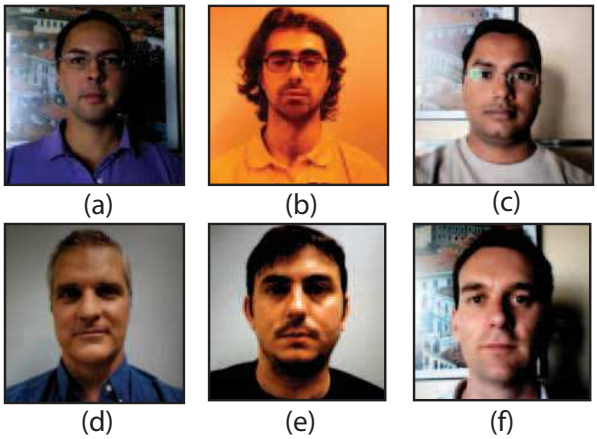

Fig. 2. Examples of face presentation attacks. A-F are examples of 2 live faces and 4 face biometric presentation attacks, where A-B are live faces, C-D are printed photo attacks and E-F are video attacks [13]. Given a still image, it is challenging to discern whether it is an image of a live authentic user or an attack. Using PPGSecure algorithm, we are able to classify with high accuracy which videos correspond to live users and which correspond to attacks.

tographs, photographs and videos on a display device, or 3-dimensional - masks. Several examples of live faces and face presentation attacks are shown in Figure 2. Each of these attacks poses different challenges for the antispoofing systems. Galbally et al. presented a review of face spoofing attacks that have been used in the past, as well as a detailed survey of attempted anti-spoofing approaches. [1].

B. Motion and Appearance Based Anti-spoofing

Prior anti-spoofing techniques can be categorized as motion-based or appearance-based [1]. Motion-based techniques considered the difference between foreground face motion and the background, or motion caused by involuntary eye movements, such as blinking [2]–[4], gaze [7] or pupillary reflex [5], [6]. In addition, differences in motion of the face in the foreground and the background have been used with optical flow [8]–[10] or motion sensors on mobile devices [11]. While effective against printed image and some video replay attacks, these motion-based techniques could not prevent attacks with high resolution 3D printed masks where eyes have been cut out [12] allowing the attacker to blink and change their gaze.

Meanwhile, some appearance-based methods used differences in texture and spectral reflectance between live faces and face presentation attacks [13]–[15], as well as differences in multispectral properties of skin and mask materials [16]. While these methods are able to distinguish between some mask attacks and a real face, they do not generalize well to new datasets and fail in cases where attackers print masks on very realistic materials [17]. Both appearance-based and motion-based methods are only able to detect a few facial recognition spoofing attacks and they fail in more challenging cases.

C. Physiology for Anti-spoofing

The idea of using physiology to prevent spoofing attacks was initially employed in fingerprint authentication [18], where a pulse oximeter was placed at the fingerprint sensor location [19] to verify that it is a real finger. There have been very few attempts to use PPG signals as a face liveness detection modality. Suh et al. used YCbCr color space and time domain PPG waveforms to distinguish between live faces and BPA [20]. Their method was only able to detect photograph attacks accurately and they used a small dataset that is not publicly available. The current state of the art method for physiology-based antispoofing is the algorithm developed by Liu et al. [17]. They computed a PPG signal using CHROM [21] method from many small facial regions and computed a similarity between each of the regions. They defined this similarity as the maximum value of the Fourier transform of the cross correlation between each two signals. The authors took into account a spatial distribution of good and poor facial regions similar to Kumar et al. [22] by putting lower weights on signals from poor regions and higher weights on good signals. Different from [22], these weights were learned through a data-driven approach from videos of a training set live subjects. Then, they trained a Support Vector Machine classifier on the weighted similarity features to classify a video of a face as a real live face or an attack.

D. Video based measurement of physiology

As the heart pumps blood through the body, the amount of blood passing through a given region of the blood vessels changes in sync with the cardiac cycle. Hemoglobin and oxyhemoglobin present in the blood absorb light most intensely in 520 - 580 nm which is within the range of the green channel spectrum in RGB cameras [23], [24]. Therefore, as blood flows, the amount of hemoglobin at a given point will change over time leading to changes in the amount of light being absorbed and causing a very small color change. Although this small color change cannot be seen with a naked eye, with careful signal processing it can be retrieved from a video recording and provide accurate vital signs measurements. Recently, there has been a rapid growth in technology for ambient light camera-based vital signs detection, such as pulse rate, pulse rate variation and breathing rate [25]. Sun and Thakor wrote a survey summarizing the current state of the art methods in PPG detection from cameras [26]. McDuff et al. found that using cyan, green, and orange (CGO) bands instead of RGB color space improves vital signs estimation [27]. To improve the signal to noise ratio in challenging scenarios, Kumar et al. used an adaptive weighted average called the goodness metric which only includes strong regions in the PPG estimate and rejects regions corrupted by noise or with very weak signals [22]. Tulyakov et al. used matrix completion to improve the PPG estimates in presence of motion by automatically selecting good facial regions [28].

E. Liveness detection Using Vital Signs

Since PPG signals detected from live skin regions share properties that differentiate them from other signals, several approaches used this property to detect liveness, or locating live skin region detection in the videos. To improve the PPG estimate, Bobbia et al. [29] used a raw PPG estimate to better locate the face region boundary by detecting the skin pulsatility in each of small regions in the face. Wang et al. [30] detected live skin regions by looking for features characteristic of PPG signals, assuming the live signals should share specific properties, such as location of the maximum frequency peak, small phase delay, small frequency spectrum entropy and large inner product. They created a matrix using these four features for each video and used an unsupervised approach of matrix factorization to find regions in the video corresponding to live regions.

Existing attempts in the literature of physiology-based anti-spoofing or liveness detection are limited to datasets with a small variety of attacks [17], [20] or do not address the more challenging issues of varying light conditions and hand motion if the camera or the form of attack is handheld [12], [17], [20]. In our proposed PPGSecure algorithm, we detect PPG signals using intensity changes observed in the green channel and use their frequency spectra directly to train a machine learning classifier, making our method more robust to diverse scenarios.

III. Methodology

A. Background: Camera-Based PPG Estimation

Flowing blood through the circulatory system causes a color change that can be observed in alive faces with a camera. When a biometric presentation attack, such as a photograph or video display, is presented to the authentication system, the captured video does not contain these subtle pulsatile color changes induced by blood flow (See Figure 3). 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 blood vessels. The amount of light absorbed by blood vessels changes with changing hemoglobin and oxyhemoglobin concentrations during the cardiac cycle. This results in a very weak time varying signal detected by the camera. On the other hand, 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, which is not sufficient to be detected by a camera.

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 face and a face BPA. We illustrate the drastic difference in the observed Fourier transforms of the obtained PPG signals from alive faces and attacks in Figure **??**.

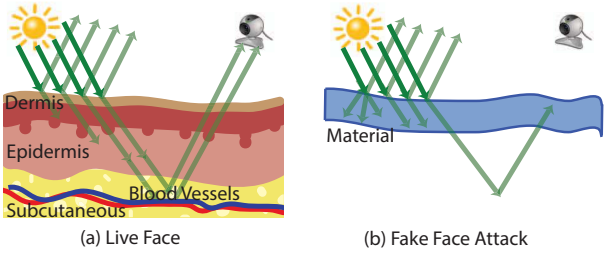

Fig. 3. PPG signals derived from color changes due to blood flow can be observed from a video recording of a live face because some of the light is able to pass through the skin and reach blood vessels. These types of color changes are not present in face attacks because there are no blood vessels present. Therefore, the observed intensity changes do not have the characteristic PPG signals properties.

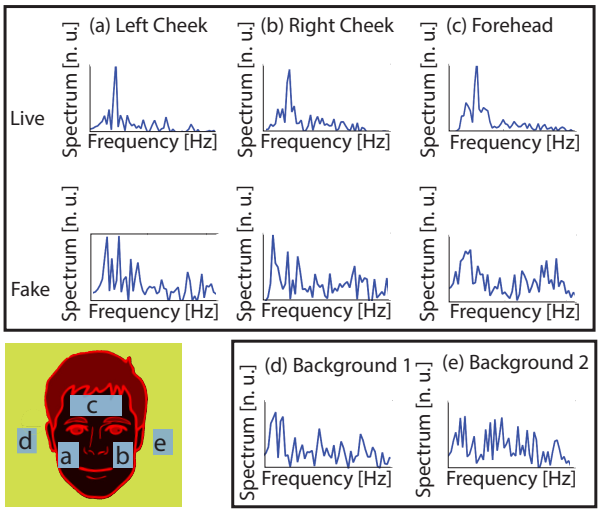

Fig. 4. PPG signals from different facial regions on a live face share characteristic similarities that are abscent in signals from a face attack or the background regions.

B. PPGSecure

Our approach is motivated by the fact that signals from different parts of a live face share similarities in their frequency spectra, while signals obtained from a presentation attack or the background will be very different from the live signals. We extracted PPG signals, computed the spectral features and used them to train a classifier

1) PPG Signal Extraction: To extract the PPG signals from the video of a facial skin region, similar to Kumar et. al. [22], we converted the RGB video to the green channel and we tracked the face using Kanade Lucas Thomasi (KLT) tracker [31]. Different from Kumar et al. , we did not compute a signal to noise ratio for each small facial region because we are not trying to improve the accuracy of vital signs estimation. Instead, we are interested in differentiating between PPG signals from a live face and noise and unrelated illumination changes. After detecting facial landmarks with

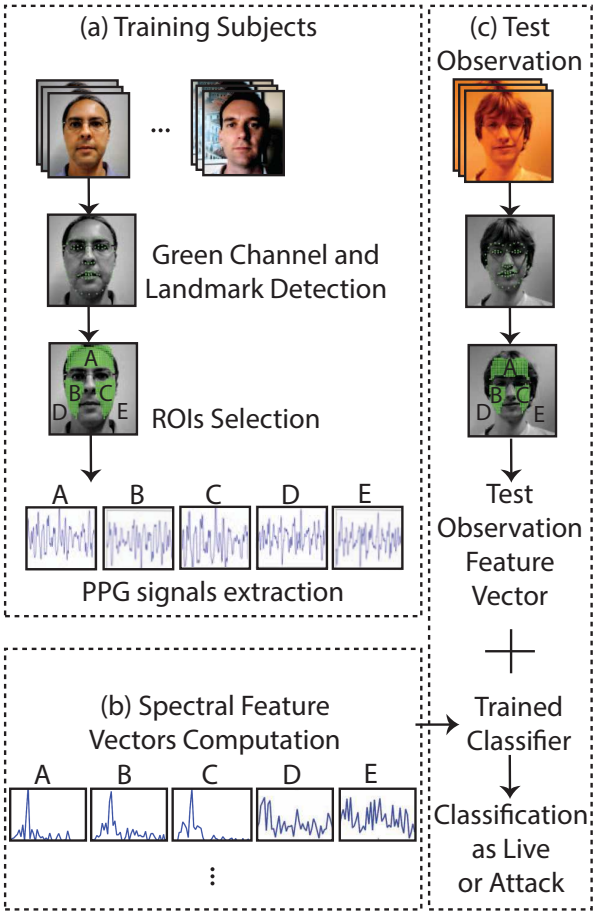

Fig. 5. Steps involved in PPGSecure. First, we detect facial landmarks and find the regions of interest (ROIs) in the face and the background. We extract PPG signals from each ROI (Part 1). We compute spectral feature vectors (Part 2) and train a machine learning classifier on training subjects' videos, which then classifies a new person's video as live or attack based on its spectral features (Part 3).

Kazemi's landmark detector [32], we selected three regions on the face known to be physiologically good for detecting PPG signals, that is the forehead, left and right cheeks. In addition to the facial regions, we selected two 50 x 50 pixels regions in the background, one to the left of the face and one to the right. The advantage of including the background regions is that any temporal variations 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. Thus adding background regions to the spectral feature vectors provides robustness against illumination fluctuations mimicking pulse signal that could fool PPGSecure. We averaged the temporal intensity changes to obtain a single PPG signal describing each region of interest. The process of extracting PPG signals is shown in Part 1 of Figure 4.

2) Spectral Features Computation: 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 [0.5 Hz, 5 Hz] range, which corresponds to physiological range of PPG signals. The magnitude of the Fourier spectrum of each filtered PPG signal becomes a spectral feature. We concatenated these spectral features from three facial regions and two background regions to obtain a spectral feature vector for classification. See Part 2 of Figure 4. Spectral features are discriminative for classifying a video of a face as live or as a face BPA because they have similarities in live faces but not in the face attacks.

3) Classification As Live Or Attack: To classify a new video of a person's face as alive or as a biometric presentation attack we used machine learning. We trained a support vector machine (SVM) [33] and a random decision forest (RDF) classifiers [34] on spectral features of training subjects' videos. We used a leave-one-subject-out validation (LOsOV) method to avoid training and testing on spectral features from videos of the same person. In LOsOV approach, the training is done on all videos in the dataset except for videos of one person. These left out videos of the same person are used as a testing set to evaluate the initial performance of the model. This procedure is repeated, each time leaving out a different person's videos and training on the remaining dataset. The final performance result is obtained by averaging the initial resutls on each individual left-out person. We trained on all kinds of attacks together (photo, video, hand-held or fixed) but we tested each attack scenario separately to understand which situations poses a greater challenge for the detection model.

IV. Results

A. Evaluation on Replay-Attack Dataset

To evaluate the performance of PPGSecure we used a publicly available dataset, Replay-Attack [13] with video and photograph biometric presentation attacks. The dataset contains 360 x 240 pixels video recordings, recorded at 25 fps with a total of 1300 videos of 50 different people. The dataset has videos of authentic live users, and video and photo presentation attacks, in controlled and adverse lighting conditions. The photo and video attacks were recorded with the form of the attack fixed and handheld in front of the camera causing small motion. We report our results separately on handheld and fixed attacks and separately on photo and video attacks.

To report our results, we used error metrics defined in terms of True Positives, True Negatives, False Positives and False Negatives, where a True Positive (TP) is an attack correctly classified as an attack, a True Negative (TN) is a live face correctly classified as live, a False Positive (FP) is a live face misclassified as an attack and a False Negative (FN) is an attack misclassified as live. The error evaluation metrics we used are defined as $Specificity = \frac{TN}{TN+FP}$, $Sensitivity = \frac{TP}{TP+FN}$, $Precision = \frac{TP}{TP+FP}$, $FalsePositiveRate\ (FPR) = \frac{FP}{FP+TP}$ and $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$. Having a higher number of attacks misclassified as live has more severe consequences than misclassifying a live face as an attack because it means that a

TABLE I

PPGSECURE PERFORMANCE ON BIOMETRIC PRESENTATION ATTACKS DETECTION USING FIXED PHOTOGRAPHS

Method	*Specificity*	*Sensitivity*	*Precision*	*FPR*	*Accuracy*
Liu [17]	99.57 %	95.28 %	99.59 %	0.41 %	97.32 %
PPGSecure	93.68 %	82.03 %	94.65 %	5.35 %	86.96 %
PPGSecure +background	96.68 %	85.97 %	97.11 %	2.89 %	90.63 %
PPGSecure +filtered	99.59 %	98.79 %	99.59 %	0.41 %	99.18 %
PPGSecure +filtered +background	**100 %**	**100 %**	**100 %**	**0 %**	**100 %**

TABLE II

PPPGSECURE PERFORMANCE ON BIOMETRIC PRESENTATION ATTACKS DETECTION USING HANDHELD PHOTOGRAPHS

Method	*Specificity*	*Sensitivity*	*Precision*	*FPR*	*Accuracy*
Liu [17]	97.40 %	81.92 %	97.91 %	2.09 %	88.15 %
PPGSecure	91.27 %	83.44 %	92.18 %	7.82 %	86.94 %
PPGSecure +background	100 %	100 %	100 %	0 %	100 %
PPGSecure +filtered	100 %	98.80 %	100 %	0 %	99.39 %
PPGSecure +filtered +background	**100 %**	**100 %**	**100 %**	**0 %**	**100 %**

TABLE III

PPGSECURE PERFORMANCE ON BIOMETRIC PRESENTATION ATTACKS DETECTION USING FIXED VIDEOS

Method	*Specificity*	*Sensitivity*	*Precision*	*FPR*	*Accuracy*
Liu [17]	96.52 %	95.14 %	96.58 %	3.42 %	95.82 %
PPGSecure	86.31 %	81.27 %	87.33 %	12.67 %	83.61 %
PPGSecure +background	98.41 %	85.70 %	98.65 %	1.35 %	91.09 %
PPGSecure +filtered	100 %	98.80 %	100 %	0 %	99.39 %
PPGSecure +filtered +background	**100 %**	**100 %**	**100 %**	**0 %**	**100 %**

TABLE IV

PPGSECURE PERFORMANCE ON BIOMETRIC PRESENTATION ATTACKS DETECTION USING HANDHELD VIDEOS

Method	*Specificity*	*Sensitivity*	*Precision*	*FPR*	*Accuracy*
Liu [17]	90.64 %	80.54 %	91.97 %	8.03 %	84.87 %
PPGSecure	100 %	84.97 %	100 %	0 %	91.16 %
PPGSecure +background	100 %	100 %	100 %	0 %	100 %
PPGSecure +filtered	100 %	98.80 %	100 %	0 %	99.39 %
PPGSecure +filtered +background	**100 %**	**100 %**	**100 %**	**0 %**	**100 %**

photograph or a video has been wrongly classified as the live authentic user and an attacker has gained access to the system. If an authentic user is incorrectly classified as an attack, it is less problematic because the user can try to access the device again.

B. Comparison to Existing Methods

We compared the performance of PPGSecure to the current state of the art physiology-based method which uses PPG signals and machine learning [17]. Liu et al.'s method is based on learning a spatial confidence map using PPG signals from several facial regions from many live people. Because they trained their algorithm on a mask dataset which is not yet publicly available, combined with a publicly available mask 3DMAD dataset [12], we are not able to directly compare our results to their performance because we only have access to a subset of their dataset. Therefore, we implemented their algorithm as described in their paper and evaluated it on the Replay-Attack public dataset [13].

PPGSecure outperforms Liu et al. on Replay-Attack dataset, especially when hand motion is present (Tables II and IV). Liu et al.'s performance drops in presence of hand motion. This could be because they look for correlated changes in the facial regions and handshake motion makes the whole photograph or video move uniformly, resulting in high cross-correlation patterns. Because hand motion frequency is different from that of live PPG signals, PPGSecure is able to learn the differences in frequency spectra patterns between handheld attacks and authentic live user's PPG signals.

PPGSecure performs better when the PPG signals are bandpass filtered before taking the Fourier transform (PPGSecure+filtered). 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 (PPGSecure+background). It is especially apparent when the signals are not bandpass filtered and contain more noise in the frequencies outside of the physiological range. This supports our hypothesis that adding the background regions reduces the error.

V. Discussion

A. Detection Accuracy Depends On Types Of Attacks

Physiology methods, such as PPGSecure and the method developed by Liu et al. detect a characteristic PPG signal pattern which is only present in live skin regions and absent in biometric presentation attacks, regardless of what face attack material was used. Some of these attacks are easier to detect with physiology methods than others. For instance, detecting a photograph attack might be easier if it is fixed on a tripod and there is no motion in the image, making the intensity changes minimal and very different from characteristic color changes present in live faces. If presented with a video recording of a PPG signal, PPGSecure should still be able to classify this as an attack. A similar challenge is present in Replay-Attack dataset when the attack is in the form of a video of a real person displayed to the authentication system on a screen and PPGSecure is able to detect these attacks with high accuracy. Future work includes the development of a liveness metric which will compute how good a given video is for detecting physiology-based liveness.

B. Difficulties Due To Quality Of Region Of Interest

Any situation where light is obstructed from reaching the blood vessels will decrease the performance of physiology-based attack detection methods. For example, when a person is wearing heavy make-up or has a darker skin tone, the accuracy of this method will be compromised. PPGSecure, as well as Liu et al. 's method, rely on using specific facial regions in live faces, where the PPG signals are physiologically strong. This approach may be problematic when people have facial hair, for instance, a beard covering the cheeks or bangs covering the forehead. When comparing the hair covered regions to other good regions, we will end up with the same facial confidence map as we would for a person wearing a mask and having a part of the mask cut out to allow PPG signals to be detected but keeping enough face covered for the facial recognition system to be fooled. Such datasets including attacks where only a part of the mask or photograph has been cut out are not available to test our assumptions but we plan to collect our own data in the future to explore this.

C. Applications Beyond Biometrics

Besides the anti-spoofing application of this work, there are other fields where being able to differentiate between live skin PPG signals from the background or noise is beneficial. Detecting liveness in videos may allow finding survivors during a rescue action by flying drones and recording the scene. Additionally, PPGSecure could facilitate detecting humans in videos, which is a challenging problem due to a wide range of poses and skin tones [35], [36]. Furthermore, realistic modeling of human skin is a difficult problem because of the complexity of biological tissues and their interaction with light [37], [38]. By considering the obtained temporal illumination changes due to the blood flow in the skin, skin models could be improved by changing the light absorption model of the skin to make their appearance more believable.

Acknowledgments

This work was supported in part by NSF grant CNS-1429047 and NHARP grant THECB-NHARP 13308.

References

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

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] K. Kollreider, H. Fronthaler, and J. Bigun, "Evaluating liveness by face images and the structure tensor," *Proceedings - Fourth IEEE Workshop on Automatic Identification Advanced Technologies, AUTO ID 2005*, vol. 2005, pp. 75–80, 2005.

[9] ——, "Non-intrusive liveness detection by face images," *Image and Vision Computing*, vol. 27, no. 3, pp. 233–244, 2009.

[10] W. Bao, H. Li, N. Li, W. Jiang, and a. O. F. Field, "A Liveness Detection Method for Face Recognition Based on Optical Flow Field," *Computer*, pp. 0–3, 2009.

[11] P. Chen, S., Pande, A., and Mohapatra, "Sensor-Assisted Facial Recognition : An Enhanced Bio- metric Authentication System for Smartphones," *In Proceedings of the 12th annual international conference on Mobile systems, applications, and services MobiSys '14*, pp. 109–122, 2014.

[12] N. Erdogmus and S. Marcel, "Spoofing in 2D face recognition with 3D masks and anti-spoofing with Kinect," *IEEE 6th International Conference on Biometrics: Theory, Applications and Systems, BTAS 2013*, 2013.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] P. D. Lapsley, J. A. Lee, D. F. Pare Jr, and N. Hoffman, "Anti-fraud biometric scanner that accurately detects blood flow," Apr. 7 1998, uS Patent 5,737,439.

[19] B. Tan, A. Lewicke, D. Yambay, and S. Schuckers, "The effect of environmental conditions and novel spoofing methods on fingerprint anti-spoofing algorithms," in *2010 IEEE International Workshop on Information Forensics and Security*. IEEE, 2010, pp. 1–6.

[20] 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.
[21] G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rppg," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886, 2013.
[22] 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.
[23] L. F. C. Martinez, G. Paez, and M. Strojnik, "Optimal wavelength selection for noncontact reflection photoplethysmography," in *International Commission for Optics (ICO 22)*. International Society for Optics and Photonics, 2011, pp. 801 191–801 191.
[24] J. A. Crowe and D. Damianou, "The wavelength dependence of the photoplethysmogram and its implication to pulse oximetry," in *Engineering in medicine and biology society, 1992 14th Annual International Conference of the IEEE*, vol. 6. IEEE, 1992, pp. 2423–2424.
[25] 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.
[26] 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.
[27] 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.
[28] 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.
[29] S. Bobbia, Y. Benezeth, and J. Dubois, "Remote Photoplethysmography Based on Implicit Living Skin Tissue Segmentation," 2016.
[30] 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.
[31] 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.
[32] 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.
[33] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
[34] 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.
[35] A. Krishnaswamy and G. V. Baranoski, "A biophysically-based spectral model of light interaction with human skin," in *Computer Graphics Forum*, vol. 23, no. 3. Wiley Online Library, 2004, pp. 331–340.
[36] G. V. Baranoski and A. Krishnaswamy, "Light interaction with human skin: from believable images to predictable models," in *ACM SIGGRAPH ASIA 2008 courses*. ACM, 2008, p. 10.
[37] Z. Lin, *Modeling shape, appearance and motion for human movement analysis*. ProQuest, 2009.
[38] X. Liang, C. Xu, X. Shen, J. Yang, S. Liu, J. Tang, L. Lin, and S. Yan, "Human parsing with contextualized convolutional neural network," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1386–1394.