

Impact of Makeup on Remote-PPG Monitoring

WENJIN WANG, AND CAIFENG SHAN

Remote Sensing Department, Philips Research, Eindhoven, The Netherlands

Abstract: Camera-based remote photoplethysmography (remote-PPG) enables contactless measurement of blood volume pulse from the human skin. Skin visibility is essential to remote-PPG as the camera needs to capture the light reflected from the skin that penetrates deep into skin tissues and carries blood pulsation information. The use of facial makeup may jeopardize this measurement by reducing the amount of light penetrating into and reflecting from the skin. In this paper, we conduct an empirical study to thoroughly investigate the impact of makeup on remote-PPG monitoring, in both the visible (RGB) and invisible (Near Infrared, NIR) lighting conditions. The experiment shows that makeup has negative influence on remote-PPG, which reduces the relative PPG strength (AC/DC) at different wavelengths and changes the normalized PPG signature across multiple wavelengths. It makes (i) the pulse-rate extraction more difficult in both the RGB and NIR, although NIR is less affected than RGB, and (ii) the blood oxygen saturation extraction in NIR impossible. To the best of our knowledge, this is the first work that systematically investigate the impact of makeup on camera-based remote-PPG monitoring.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

OCIS codes: (280.0280) Remote sensing and sensors; (170.3880) Medical and biological imaging.

References and links

1. W. Verkruyse, L. Svaasd, and J. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Exp.* **16**(26), 21 434–21 445 (2008).
2. Y. Sun and N. Thakor, "Photoplethysmography revisited: from contact to noncontact, from point to imaging," *IEEE Trans. Biomed. Eng.* **63**(3), 463–477 (2016).
3. D. J. McDuff, J. R. Estepp, A. M. Piasecki, and E. B. Blackford, "A survey of remote optical photoplethysmographic imaging methods," in *Proceedings of IEEE conference on Engineering in Medicine and Biology Society* (IEEE, 2015), pp. 6398–6404.
4. W. Wang, "Robust and automatic remote photoplethysmography," Technische Universiteit Eindhoven, D. Phil. Thesis (2017).
5. A. A. Kamshilin, E. Nippolainen, I. S. Sidorov, P. V. Vasilev, N. P. Erofeev, N. P. Podolian, and R. V. Romashko, "A new look at the essence of the imaging photoplethysmography," *Scientific reports* , **5**, p. 10494 (2015).
6. A. V. Moco, S. Stuijk, and G. de Haan, "New insights into the origin of remote PPG signals in visible light and infrared," *Scientific reports* **8**(1), p. 8501 (2018).
7. W. Wang, AC. den Brinker, S. Stuijk, and G. de Haan, "Robust heart rate from fitness videos," *Physiol. Meas.* **38**(6), p. 1023 (2017).
8. S. Leonhardt, L. Leicht, and D. Teichmann, "Unobtrusive vital sign monitoring in automotive environmentsa review," *Sensors* **18**(9), p. 3080 (2018).
9. S. Liu, X. Lan, and P. C. Yuen, "Remote photoplethysmography correspondence feature for 3d mask face presentation attack detection," in *Proceedings of the European Conference on Computer Vision (ECCV, 2018)*, pp. 558–573.
10. J. Hernandez-Ortega, J. Fierrez, A. Morales, and P. Tome, "Time analysis of pulse-based face anti-spoofing in visible and NIR," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW, 2018)*, pp. 544–552.
11. G. Heusch, and S. Marcel, "Remote blood pulse analysis for face presentation attack detection," In *Handbook of Biometric Anti-Spoofing*, pp. 267–289 (2019).
12. A. Prati, C. Shan, and K. I. K. Wang, "Sensors, vision and networks: From video surveillance to activity recognition and health monitoring," *Journal of Ambient Intelligence and Smart Environments* **11**(1), 5–22 (2019).
13. W. Wang, S. Stuijk, and G. de Haan, "Exploiting spatial redundancy of image sensor for motion robust rPPG," *IEEE Trans. Biomed. Eng.* **62**(2), 415–425 (2015).
14. M. Lewandowska, J. Ruminski, T. Kocejko, and J. Nowak, "Measuring pulse rate with a webcam - a non-contact method for evaluating cardiac activity," in *Proceedings of Federated Conference on Computer Science and Information Systems (FedCSIS, 2011)*, pp. 405–410.
15. M. Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.* **58**(1), 7–11 (2011).

16. G. de Haan and V. Jeanne, “Robust pulse rate from chrominance-based rPPG,” *IEEE Trans. Biomed. Eng.* **60**(10), 2878–2886 (2013).
17. G. de Haan and A. van Leest, “Improved motion robustness of remote-PPG by using the blood volume pulse signature,” *Physiol. Meas.* **35**(9), 1913–1922 (2014).
18. W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, “Algorithmic principles of remote-PPG,” *IEEE Trans. Biomed. Eng.* **64**(7), 1479–1491 (2017).
19. W. Wang, S. Stuijk, and G. de Haan, “A novel algorithm for remote photoplethysmography: Spatial subspace rotation,” *IEEE Trans. Biomed. Eng.* **63**(9), 1974–1984 (2016).
20. D. McDuff, and E. Blackford, “iPhys: An open non-contact imaging-based physiological measurement toolbox. arXiv preprint arXiv:1901.04366 (2019).
21. A.M. Unakafov, “Pulse rate estimation using imaging photoplethysmography: generic framework and comparison of methods on a publicly available dataset,” *Biomed. Physics and Eng. Exp.* **4**(4), pp. 045001 (2018).
22. Q. Zhang, Y. Zhou, S. Song, G. Liang, and H. Ni, “Heart rate extraction based on near-infrared camera: Towards driver state monitoring,” *IEEE Access* **6**, pp. 33076–33087 (2018).
23. G. Balakrishnan, F. Durand, and J. Guttag, “Detecting pulse from head motions in video,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR, 2013)*, pp. 3430–3437.
24. T. Vogels, M. van Gastel, W. Wang, and G. de Haan, “Fully-automatic camera-based pulse-oximetry during sleep,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW, 2018)*, pp. 1349–1357.
25. W. Wang, B. den Brinker, and G. de Haan, “Single element remote-PPG,” *IEEE Trans. Biomed. Eng.* **99**(0), pp. 1–1 (2019). (accepted, DOI: 10.1109/TBME.2018.2882396)
26. M. van Gastel, S. Stuijk, and G. de Haan, “New principle for measuring arterial blood oxygenation, enabling motion-robust remote monitoring,” *Scientific reports* **6**, p. 38609 (2016).
27. A. Zadeh, Y. Chong Lim, T. Baltrušaitis, and L. P. Morency, “Convolutional experts constrained local model for 3d facial landmark detection,” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV 2017)*, pp. 2519–2528.
28. D. Shao, T. Yang, C. Liu, F. Tsow, H. Yu, and N. Tao, “Noncontact monitoring breathing pattern, exhalation flow rate and pulse transit time,” *IEEE Trans. Biomed. Eng.* **61**(11), pp. 2760–2767 (2014).
29. A. V. Moco, L. Z. Mondragon, W. Wang, S. Stuijk, and G. de Haan, “Camera-based assessment of arterial stiffness and wave reflection parameters from neck micro-motion,” *Physiological measurement* **38**(8), p. 1576 (2017).
30. A. A. Kamshilin, S. Miridonov, V. Teplov, R. Saarenheimo, and E. Nippolainen, “Photoplethysmographic imaging of high spatial resolution,” *Biomed. optics exp.*, **2**(4), 996–1006 (2011).
31. W. Wang, S. Stuijk, and G. de Haan, “Unsupervised subject detection via remote PPG,” *IEEE Trans. Biomed. Eng.* **62**(11), 2629–2637 (2015).
32. W. Wang, S. Stuijk, G. de Haan, “Living-skin classification via remote-PPG,” *IEEE Trans. Biomed. Eng.* **64**(12), 2781–2792 (2017).

1. Introduction

Camera-based remote photoplethysmography (remote-PPG) enables contactless measurement of human cardiac activities by detecting the pulse-induced subtle color changes from human skin surface [1]. It can be used for various healthcare applications (e.g. patient monitoring, neonate monitoring, sleep monitoring, general ward, emergency department triage, cardio-fitness training, driver monitoring in automotive, etc.) by measuring different vital signs (e.g. pulse rate (variability), respiratory rate, blood oxygen saturation, pulse transit time, etc.) from human face and body [2, 3].

The fundamental assumption on the optical-physiological operation of remote-PPG is that the light can penetrate into the skin, travel through relatively deep vasculature (e.g., arterioles), and “see” blood absorption variations in the arteries [4]. And a camera can receive the amount of light reflected from the skin that carries the information of blood volume changes. In this procedure, an essential requirement is that light can travel back and forth between different layers of the skin microvascular bed and reach deeper arteries containing pulsatile blood. There is a controversy/debate [5, 6] on whether the PPG signal measured by a contactless camera only stems from the blood absorption variations in deep arterioles, or also from the elastic deformations of the capillary bed due to the pulse oscillations. There is less confusion for the Near Infrared (NIR) wavelength sensing as the light in this range can penetrate deeper into the skin, but it could be a consideration for the RGB wavelength sensing (i.e. whether blue/green wavelength can equally probe arterioles).

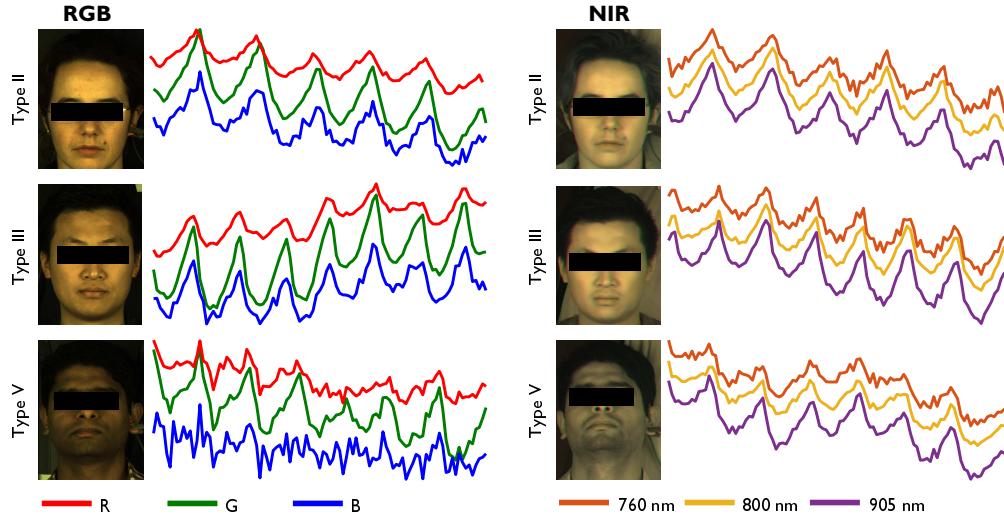


Fig. 1. The raw wavelength signals obtained from three subjects with different skin types in RGB and NIR. It shows that the blue wavelength signal measured from skin-type V (dark skin) contains little blood volume pulse information, due to strong absorption of the skin melanin at this wavelength.

The influence of skin property on remote-PPG, especially the skin melanin (skin-color related), has been studied [13] in terms of the skin types according to the Fitzpatrick scale. Skin melanin is a deteriorating factor for PPG extraction in RGB, but less problematic in NIR (see Fig. 1). The reason is: (i) the melanin content (e.g. eumelanin and pheomelanin monomers) has higher absorption of light in RGB; (ii) the NIR light can penetrate deeper into the skin. This is especially challenging for the blue wavelength signal measured from a subject with darker skin (e.g. skin-type V), which may contain little blood volume pulse information (see Fig. 1), i.e. blue light is strongly absorbed by the melanin and it has shallow penetration depth into the skin. If light cannot travel deep into the skin, the diffuse reflection signal measured at this wavelength will contain little PPG information. The observation on the skin melanin triggers our curiosity/question on the other skin-related factor - makeup.

The skin makeup, similar to a light filter layer, will physically reduce the amount of light penetrating into and reflecting from the skin, disrupting the remote measurement of the PPG signal. Thus, the influence of makeup needs to be critically considered for the application/customization of remote-PPG technology. This is because that (i) the subject face is the most commonly used region of interest (RoI) for camera-based measurement, which is also the typical skin area for the use of makeup; (ii) in some non-clinical/non-patient application scenarios, subjects may have very frequent use of makeup, such as cardio training in fitness exercise [7], driver monitoring in automotive [8], living-skin based anti-spoofing [9–11], video surveillance [12], etc.

The methodologies for remote-PPG have been advanced rapidly in recent year, particularly on addressing the challenges of body motion and ambient light changes [13–19]. Algorithmic solutions for remote-PPG (i.e. extraction of the PPG signal from camera wavelength signals) also become matured/standardized for reproducible study [20, 21]. However, there is so far no study that thoroughly investigate the influence of makeup, even though it is an obvious challenge (or detrimental factor), i.e. it was briefly mentioned in earlier works [22] focused on other applications; but has not been critically evaluated. Therefore, in this paper, we perform a study to thoroughly investigate the impact of makeup on remote-PPG, in both the visible (RGB) and invisible (NIR) lighting conditions. These are two equally important monitoring conditions/setups for the application of remote-PPG, i.e. RGB for fitness and NIR for automotive.

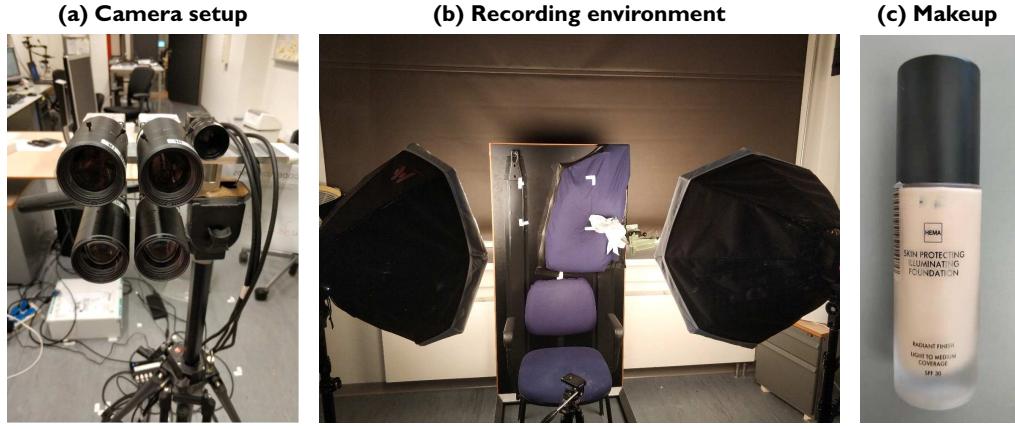


Fig. 2. Illustration of the camera setup (including multi-wavelength RGB and NIR cameras), recording environment, and facial makeup used for this study.

Our hypothesis is that makeup has negative influence on the camera-based PPG measurement, including all physiological variables derived from the measured PPG signal, such as pulse rate, pulse-rate variability, blood oxygen saturation. Given a multi-wavelength camera setup (RGB or/and NIR), we expect that the degradation caused by makeup are from twofold: (i) the relative pulsatile strength (AC/DC) will be reduced for all wavelengths, leading to lower Signal-to-Noise Ratio (SNR) of the measured PPG signal; (ii) the normalized PPG signature across multiple wavelengths (also known as the blood volume pulse signature [17]) will be changed as the amount of reduced pulsatile strength may be different at different wavelengths. These together will make the PPG-based vital signs monitoring more difficult: (i) the signal with lower relative PPG strength is more susceptible to distortions (e.g. body motion and lighting challenges), (ii) makeup-induced pulsatile signature changes will fail the signature-/knowledge-based multi-wavelength PPG extraction, and also the measurement of blood oxygen saturation that relies on the relative amplitude changes between two different wavelengths (e.g. red and infrared). We also assume that makeup has less influence on NIR as compared to RGB, because near infrared light (in the longer wavelength range) has deeper skin interrogation depth and thus is more immune to this challenge. To design a set of experiments to validate our hypothesis, our research questions are set out to be:

- *Whether make-up has influence on remote-PPG monitoring?*
- *How does it influence the PPG measurement in RGB and NIR?*
- *Is it still possible to measure pulse rate and blood oxygen saturation when skin has makeup?*

We mention that the motion-based ballistocardiograph (BCG) measurement [23] does not exploit the principle of blood absorption variations of the skin, which will not be affected by makeup and thus is excluded from our study.

2. Materials and devices

In this section, we first introduce the experimental setup. Next, we present the recording protocols. Finally, we discuss the benchmark criteria used for analyzing the impact of makeup. Unless stated otherwise, we use the following mathematical conventions throughout the paper. Italic characters denote values. Boldface characters denote vectors and matrices.

2.1. Experimental setup

As the goal of this study is to investigate the impact of makeup on remote-PPG in the visible and invisible lighting conditions, we construct an video recording setup that include both the RGB

and NIR cameras. The recording setup and experimental environment are shown in Fig. 2 (a)-(b). Different components in the setup will be introduced separately in the following text.

(i) **Cameras** The camera setup includes a 3-wavelength RGB camera and a 3-wavelength NIR camera. Both are synchronized for video recording:

- For the RGB camera, we use the global shutter RGB camera USB UI-2230SE-C from IDS (Sony ICX204, CCD sensor), with 640x480 pixels, 8 bit depth, and 15 frames per second (fps) recording rate. The camera lens is Tamron M12VM412.
- For the NIR camera, we use three separate monochrome cameras sampled at different NIR wavelengths to mimic a multi-spectral measurement. Three monochrome cameras are centered at 760 nm (bandwidth 20 nm), 800 nm (bandwidth 20 nm), and 905 nm (bandwidth 33 nm) by using NIR optical filters. The monochrome camera type is Global shutter Manta G-283 of Allied Vision Technologies GmbH (Sony ICX674ALG, CCD sensor), with 968x728 pixels, 8 bit depth, and 15 fps recording rate. The camera lens is Schneider-Kreuznach Tele-Xenar 1:4/150.

We mention that 15 fps recording rate meets the requirement of this study [24], as we use the frequency spectrogram of a pulse signal to investigate its quality, i.e. the measure of average pulse rate. The cameras are placed around 2 m in front of the subject. With the used focal length, it results in approximately 30-40% skin area in a video frame. All auto-adjustment functions (e.g. auto-focus, auto-gain, auto-white-balance, auto-exposure) of the camera are turned off during the recording. All videos are recorded in an uncompressed video format at constant frame rate.

(ii) **Light source** The illumination sources are two incandescent light fixtures consisting of 9 lamps each. Each fixture is powered at $220\text{ V} \times 1.2\text{ A} = 264\text{ W}$, providing sufficient energy and diffuse illumination for the visible and invisible sensing. Note that there is no other light source involved, i.e. daylight is prohibited.

(iii) **Subjects and makeup** A total of 22 healthy adult subjects (aged between 25 and 45), with different skin tones categorized from type-I to type-V according to the Fitzpatrick scale, participate in the experiment. This study has been approved by the Internal Committee Biomedical Experiments of Philips Research, and informed consent has been obtained from each subject. The makeup used for the experiment is skin protecting illuminating foundation Rose 01 (see Fig. 2 (c)). Fig. 3 shows the wavelength spectra of the skin with/without the makeup. Since the used makeup has a skin-similar color, the shape of the skin reflectance spectra does not vary much but its brightness is increased after applying the makeup.

Since the actual use of makeup in real-life is subject-/culture-dependent (e.g. some prefer thicker makeup, while some prefer smoother/lighter makeup with more natural look), it would be difficult to quantify all individual cases in our experiments. To this end, we aim at investigating the challenging use-case with relatively thicker makeup as a boundary condition. Critically speaking, if the makeup jeopardizes remote-PPG in certain (challenging/boundary) conditions, it will be an issue for practical usage, because we cannot foresee which/how makeup will be used by the subject in the wild, i.e. in this paper the challenging situation is considered.

2.2. Recording protocols

During the video recording, each subject is instructed to stay stationary with his/her head supported by a tripod to eliminate the BCG motion. This is to make PPG as the single source of origin. Each recording has two parts. In the first part (60 seconds), the subject has no makeup on face. In the second part (60 seconds), we put makeup on the subject forehead. The subject remains stationary throughout the recording. The reason of selecting the forehead skin area for the comparative study is that it has stronger skin pulsatility [1] that makes the analysis of PPG extraction easier (more intuitive). We also mention that the reason of keeping the recording short

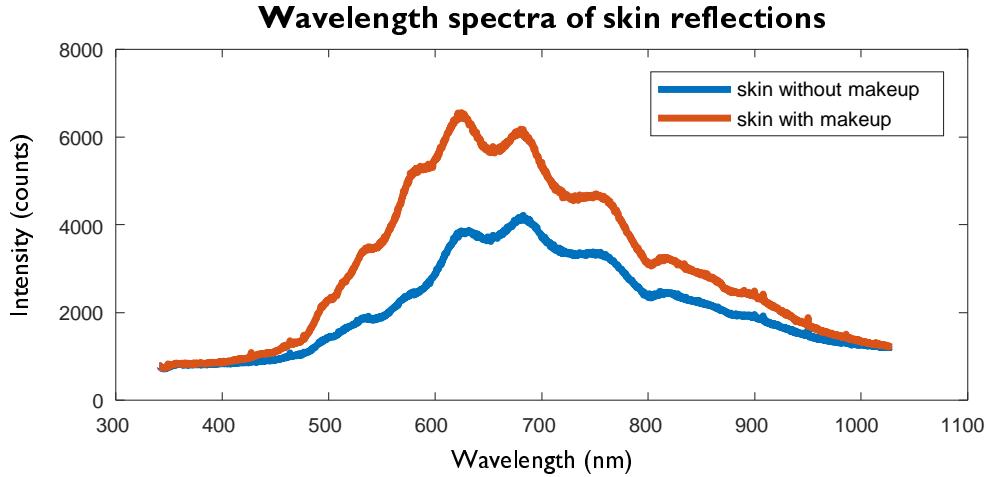


Fig. 3. The wavelength spectra of the skin without/with makeup (type: skin protecting illuminating foundation Rose 01), which is measured by reflectance spectrometer. The incandescent lamp is used as the light source to provide energy for visible and invisible wavelength sensing.

(2 minutes in total for each subject) is to avoid possible physiological changes that may occur in the longer-term recording, as it aims at comparing the skin pulsation change that is due to makeup but not other physiological reasons.

Fig. 4 exemplifies the snapshots of the RGB and NIR recordings, which show a subject without and with makeup on her forehead, i.e. the NIR images are created by plugging the three NIR channels (760 nm, 800 nm and 905 nm) into an RGB rendering system. Since the forehead area is used as the skin ROI in our study, we use a facial landmark tracker [27] to automatically locate the forehead skin area for extracting the wavelength signals.

2.3. Benchmark criteria

In order to quantify the skin pulsatility changes before and after using makeup, we use two different criteria to evaluate the wavelength signals measured from the skin:

- The *relative PPG strength (AC/DC)* at different wavelengths, which is a direct indicator for the skin pulsatility. In our approach, we first derive the DC-normalized signals (with zero mean) based on the raw input wavelength signals. This step can be achieved by the temporal normalization:

$$\tilde{\mathbf{C}}_i = \frac{\mathbf{C}_i}{\mu(\mathbf{C}_i)} - 1, \quad (1)$$

where \mathbf{C}_i and $\tilde{\mathbf{C}}_i$ denotes the i -th wavelength signal and its DC-normalized version; $\mu(\cdot)$ denotes the averaging operator. After that, $\tilde{\mathbf{C}}_i$ is pre-processed by a 4-th order Butterworth band-pass zero-phase filter with cutoff frequencies [0.6, 3] Hz for rejecting apparent distortions outside the heart-rate band. Next, we measure the median peak-to-valley distance of the DC-normalized signal ($\tilde{\mathbf{C}}_i$) as its relative PPG strength (AC/DC):

$$A_i = \text{median}(\text{peak}(\tilde{\mathbf{C}}_i)) - \text{median}(\text{valley}(\tilde{\mathbf{C}}_i)), \quad (2)$$

where **peak**(\cdot) and **valley**(\cdot) detect the peaks and valleys of a signal, respectively; **median**(\cdot) takes the median value of a vector. More specifically, we use the *findpeaks* function in Matlab R2017b to detect the peaks and valleys in a signal. The relative PPG strength can be derived from a short signal interval (by using a temporal sliding window) and

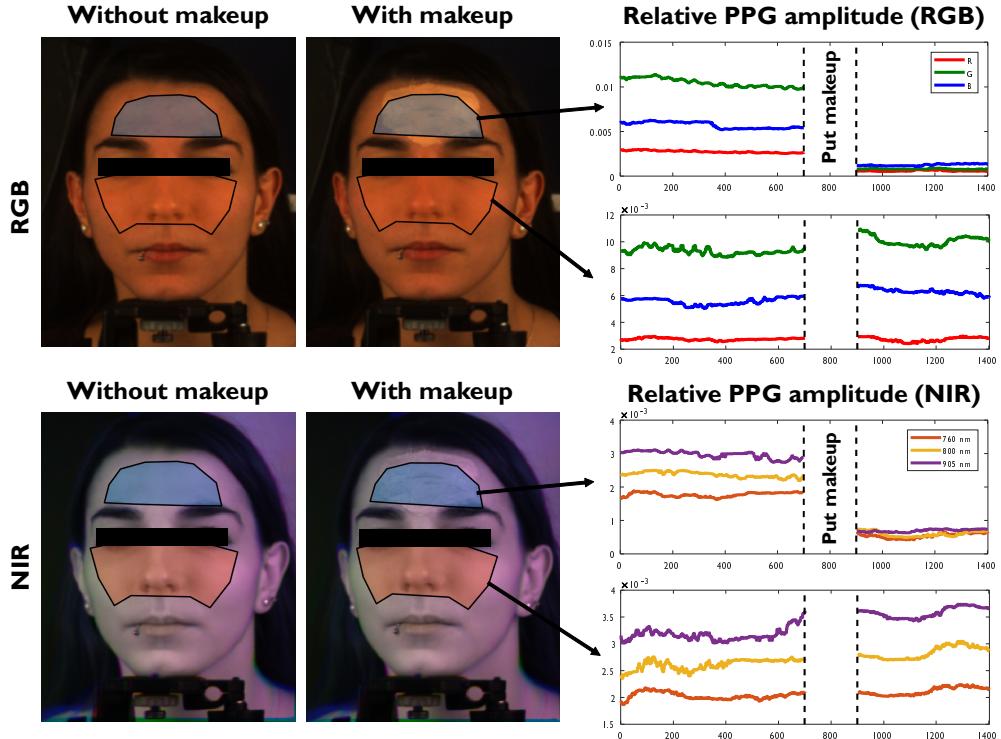


Fig. 4. Comparison of relative PPG strengths (AC/DC) at two different skin areas: forehead with makeup interference (blue region of interest) and cheek without makeup interference (red region of interest), i.e. cheek is used as the reference for visualizing the physiology-induced relative PPG amplitude changes.

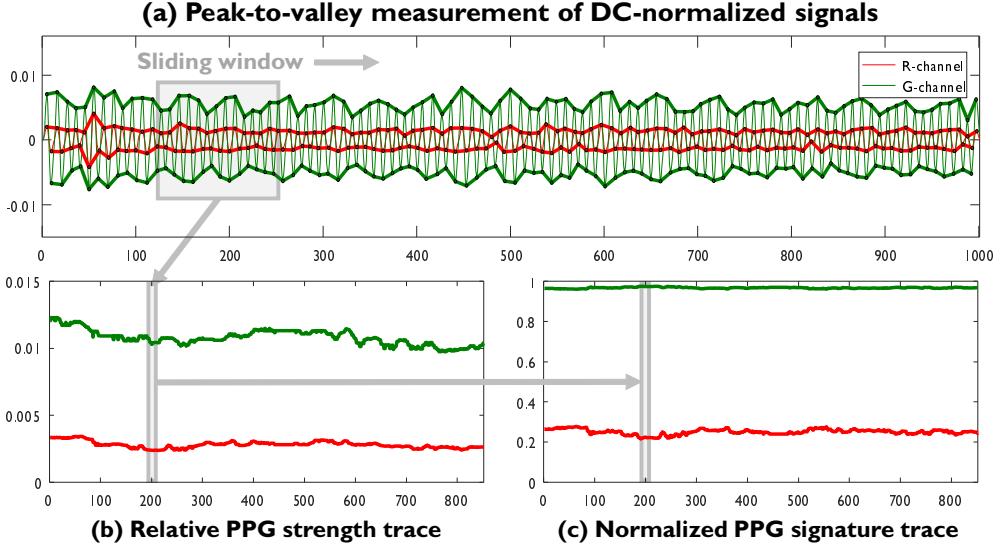


Fig. 5. Illustration of using (a) peak-to-valley measurement of the DC-normalized signals to generate (b) the relative PPG strength trace and (c) the normalized PPG signature trace, where the R and G channels are exemplified. The sliding window length is 150 frames (corresponding to 10 s for a 15 fps camera) with the sliding step 1 frame.

concatenated in time to generate a relative PPG strength trace. This procedure is illustrated in Fig. 5. If the relative PPG strength is reduced after using makeup, the performance of PPG extraction will drop.

- The *normalized PPG signature* across multiple wavelengths, which is also known as the blood volume pulse vector [17] that has been used for measuring the pulse rate [25] and blood oxygen saturation [26]. The normalized PPG signature is defined as the a unit-length vector (i.e. L2-norm is 1) containing relative PPG strengths of multiple wavelengths. Each entry of the normalized PPG signature is calculated as:

$$\bar{\mathbf{A}}_i = \frac{\mathbf{A}_i}{\sqrt{\sum \mathbf{A}_i^2}}. \quad (3)$$

Similarly, by using a temporal sliding window, we can generate a temporal signature trace of $\bar{\mathbf{A}}$ (see Fig. 5). If the signature is changed after using makeup, the signature-based (or prior knowledge based) PPG extraction will be less accurate, and the blood oxygen saturation extraction will be impossible, as it requires accurate measurement of relative PPG amplitude changes between at least two different wavelength channels (e.g. red and infrared).

To further verify whether the skin pulsatility changes in our experiment are mainly due to the use of makeup rather than other physiological reasons, we compare the pulsatility changes in other skin areas (e.g. cheek) that do not apply makeup during the recording. Fig. 4 exemplifies a comparison of relative PPG strength (AC/DC) measured from both the forehead and cheek area. It shows that the relative PPG strength measured from the cheek is relatively stable in RGB and NIR (i.e. no major variation), whereas for the forehead it is significantly reduced after using makeup. We confirm that during the 2-minutes recording, the test subject remains stationary with head supported by a tripod and has no other activity that can significantly change the physiological state, such as pulse rate (variability), respiratory rate, blood pressure, blood oxygen saturation, blood glucose, etc.

Given the wavelength signals (RGB and NIR) measured from the forehead area of the subjects, we use four multi-wavelength PPG extraction algorithms, PCA [14], ICA [15], POS [18] and PBV [17], to extract the PPG signal. The essence of multi-wavelength PPG extraction is using multi-channel combination to suppress non-pulsatile components (distortions) and emphasize the pulsatile component. These four algorithms have been widely used (or thoroughly benchmarked) in the literature [18] in terms of the PPG-signal extraction. The purpose of this study is not on comparing the performance between these four algorithms but investigating the feasibility of measuring PPG from the skin with makeup. All methods have been implemented in Matlab R2017b (using Signal Processing Toolbox) and run on a laptop with an Intel Core i7 processor (2.70 GHz) and 8 GB RAM. The parameters/configurations of these methods are remained the same as described in their original publications. We note that since POS and PBV are prior knowledge based approaches, the projection-axes of POS and the blood volume pulse signature of PBV are adapted for the processing in NIR. More specifically, the projection axes used by POS is $[0, 1, -1; -2, 1, 1]$ for RGB (considering the order of R - G - B) and $[0, -1, 1; -2, 1, 1]$ for NIR (considering the order of 770 nm - 800 nm - 905 nm). The blood volume pulse signature used by PBV is $[0.33, 0.77, 0.55]$ for RGB and $[0.46, 0.57, 0.68]$ for NIR.

3. Results and discussion

In this section, we present the experimental results and discuss the observations. Fig. 6 shows the relative PPG strengths (AC/DC) and normalized PPG signatures obtained from each individual subject, measured by the RGB and NIR camera in conditions without/with makeup. It is clear that

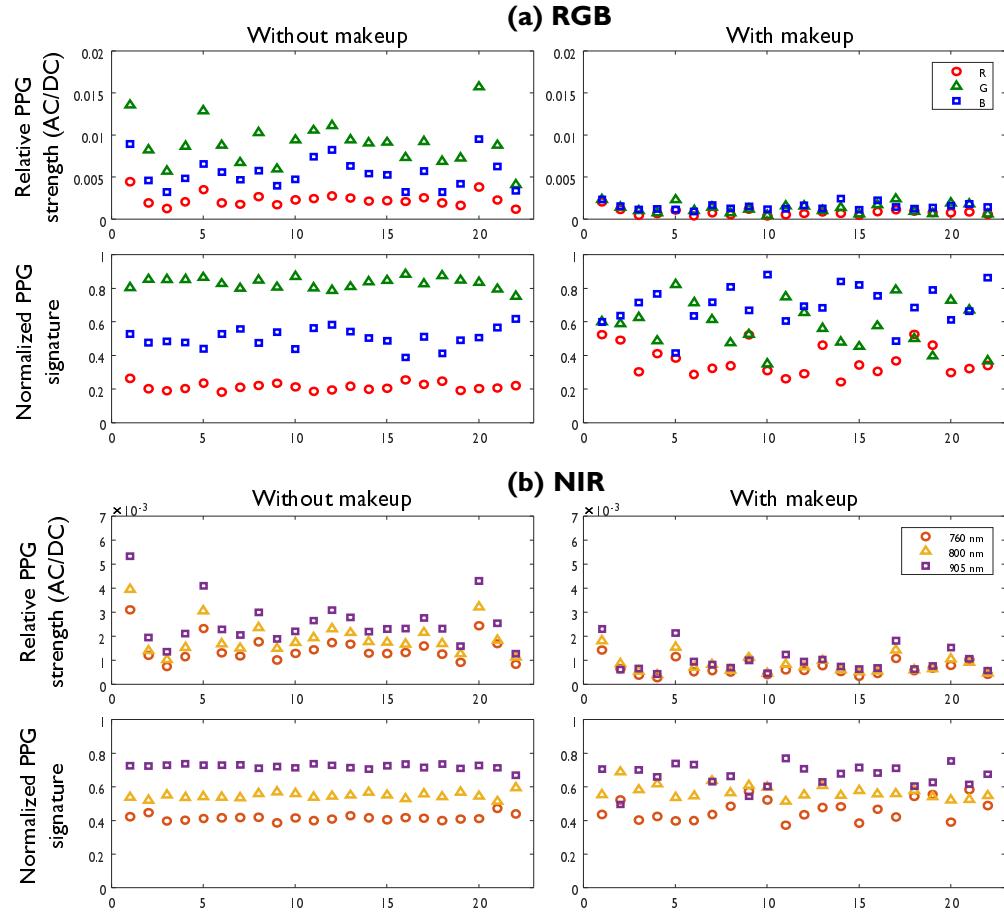


Fig. 6. The relative PPG strength (AC/DC) and normalized PPG signature of each individual test subject, obtained by the RGB and NIR cameras in conditions without and with makeup.

after using makeup, the relative PPG strengths are reduced for all wavelength settings, and the normalized PPG signatures are also modified. More specifically, the normalized PPG signatures in RGB shows a clear pattern of $G > B > R$ without using makeup. Such a characteristic channel ranking is due to the blood absorption and light penetration depth into the skin [18]. This pattern disappears due to the use of makeup. It also holds for NIR that the characteristic channel ranking of $905 \text{ nm} > 800 \text{ nm} > 760 \text{ nm}$ is interfered by makeup, though it is somehow still more visible than that in RGB.

Fig. 7 shows statistical plots of the results in Fig. 6. It confirms our observation that in both the RGB and NIR, (i) makeup clearly reduces the relative PPG strengths (AC/DC), while the pulsatility reduction is less dramatic in NIR; (ii) makeup modifies the normalized PPG signatures, but the NIR part is less influenced. For example (see Fig. 7 (b)), the characteristic channel ranking of the signature for RGB is changed from $G > B > R$ to $B > G > R$, whereas for NIR it remains the unchanged (e.g. $905 \text{ nm} > 800 \text{ nm} > 760 \text{ nm}$), i.e. only the relative contrast/ratio between the NIR channels has been changed. In order to quantify the changes caused by makeup, we show the averaged relative PPG strengths in Table I and the averaged normalized PPG signatures in Table II. As can be seen, the percentage of reduction of the relative PPG strength is 64.4%-85.2% for RGB and 54.5%-61.4% for NIR. The PPG signature is, on average, rotated 20.4° in the normalized RGB space and 4.4° in the normalized NIR space. These together suggest that makeup indeed has

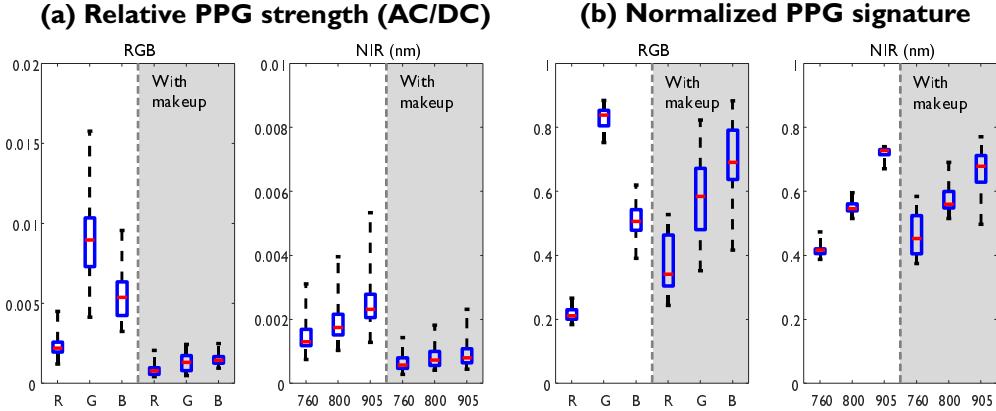


Fig. 7. The statistical plots of the relative PPG strengths (AC/DC) and normalized PPG signatures over 22 test subjects, obtained in RGB and NIR without/with makeup. The median values are indicated by horizontal bars inside the boxes, the quartile range by blue boxes, the full range by black whiskers.

Table 1. The average relative PPG strengths measured over 22 subjects (in RGB and NIR), with and without makeup.

Relative PPG strength	RGB			NIR		
	R	G	B	760 nm	800 nm	905 nm
Without makeup	0.0024	0.0091	0.0055	0.0015	0.0019	0.0026
With makeup	0.0008	0.0013	0.0015	0.0007	0.0008	0.0010
Percentage of reduction	64.4%	85.2%	72.4%	54.5%	57.2%	61.4%

less influence in NIR monitoring, i.e. it leads to less PPG-strength reductions and less signature changes.

The decreased relative PPG strength will make the PPG extraction more difficult, as the SNR of the wavelength signals become lower. The modified PPG signature is a typical threat for prior knowledge based PPG extraction methods (e.g. POS and PBV). To understand the actual effect of makeup on pulse extraction, we use aforementioned four remote-PPG algorithms to extract the PPG signals from both the clean and makeup-polluted measurement. The pulse spectrograms obtained by PCA, ICA, POS and PBV are shown in Figs. 10-13 in the Appendix. In RGB, only 8 subjects out of 22 subjects (denoted as 8/22) are feasible for PPG extraction when applying makeup on the forehead. This percentage is higher for NIR (13/22), but still approximately half of measurement fail. Though makeup seems to have less impact on NIR for pulse-rate measurement, it is still not possible to measure the blood oxygen saturation (in NIR) from the makeup-polluted skin. This is because that the relative PPG amplitude between different wavelength channels, e.g. the R channel (760 nm) and NIR channel (905 nm), has been modified (as indicated by the rotated signature), which is the basis for calculating/calibrating the oxygen saturation values. Although focusing on the discussion of pulse rate (based on PPG frequency) and blood oxygen saturation (based on PPG amplitude), we mention that makeup will affect other PPG-based measurement as well, such as pulse transit time (based on PPG phase) [28] and pulse wave velocity (based on PPG waveform morphology) [29].

We also expect that the PPG-imaging [30] based living-skin anti-spoofing will break down due to the use of makeup. To verify this, we select a test subject, whose pulse rate can be

Table 2. The average normalized PPG signatures measured over 22 subjects (in RGB and NIR), with and without makeup.

Normalized PPG signature	RGB	NIR
Without makeup	[0.22 0.83 0.51]	[0.42 0.55 0.72]
With makeup	[0.38 0.59 0.71]	[0.47 0.58 0.67]
Degree of rotation	20.4°	4.4°

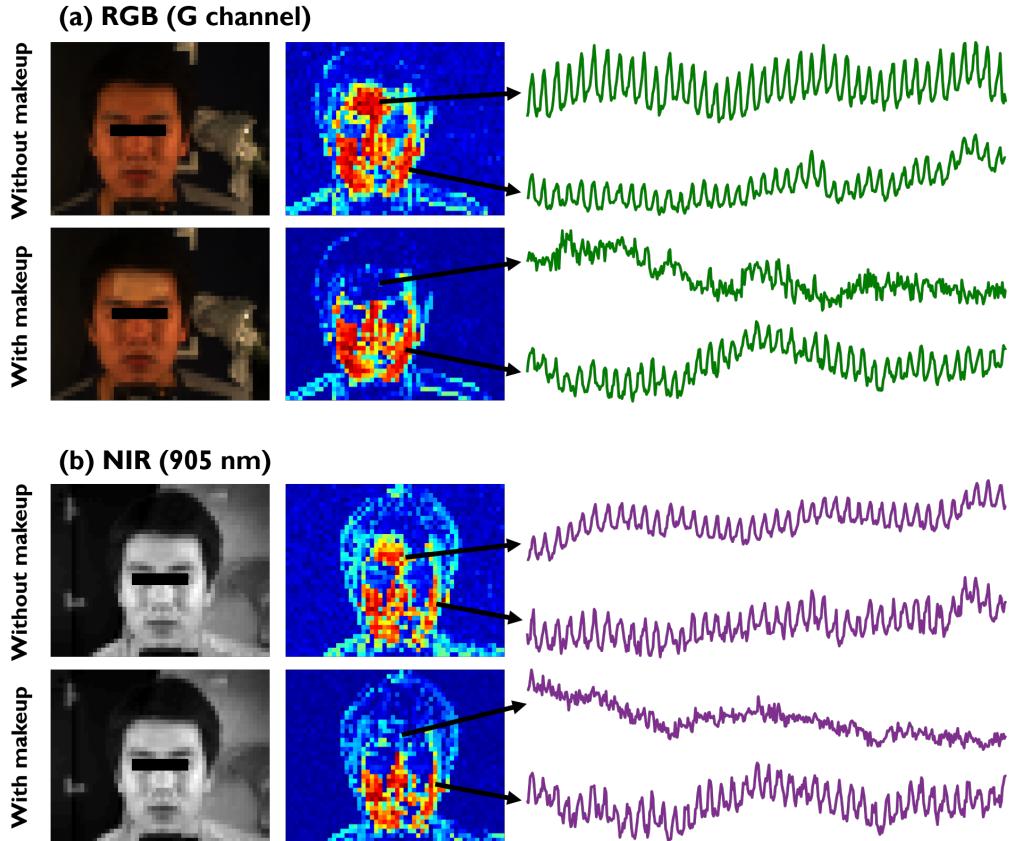
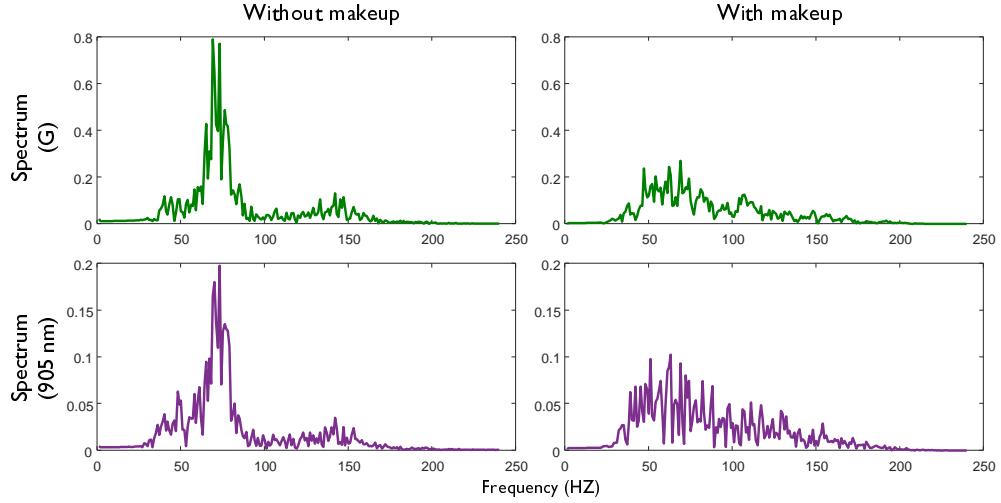


Fig. 8. The PPG imagers, i.e. 2D maps consists of spatial distribution of relative PPG amplitudes (AC/DC), of a test subject obtained in (a) RGB and (b) NIR, where forehead skin area has makeup interference. The video frames are block-wise (e.g. averaging spatial pixels) downsampled to 10% of the original frame resolution for reducing quantization noise. For RGB, the G channel is used for illustration. For NIR, the 905 nm channel is used. The raw wavelength signals at different skin sites (forehead and cheek) are shown for comparison.

measured in RGB and NIR when applying makeup on his forehead, for generating the PPG imagers (see Fig. 8). It shows that the subject's forehead area is much less pulsatile after applying the makeup, i.e. the raw signals of the G channel (in RGB) and 905 nm channel (in NIR) become much noisier, where the PPG component is not visible. The reason why we can measure the subject's pulse rate in Fig. 8 using remote-PPG algorithms (in Appendix) but cannot see clear pulsatile components in PPG imagers is: the wavelength signals used for remote-PPG extraction (in Appendix) is averaged from larger forehead area where quantization noise is reduced. The wavelength signals used for PPG imaging are from higher resolution spatial blocks that have

(a) Spectrum (1 subject)



(b) Spectrum skewness (22 subjects)

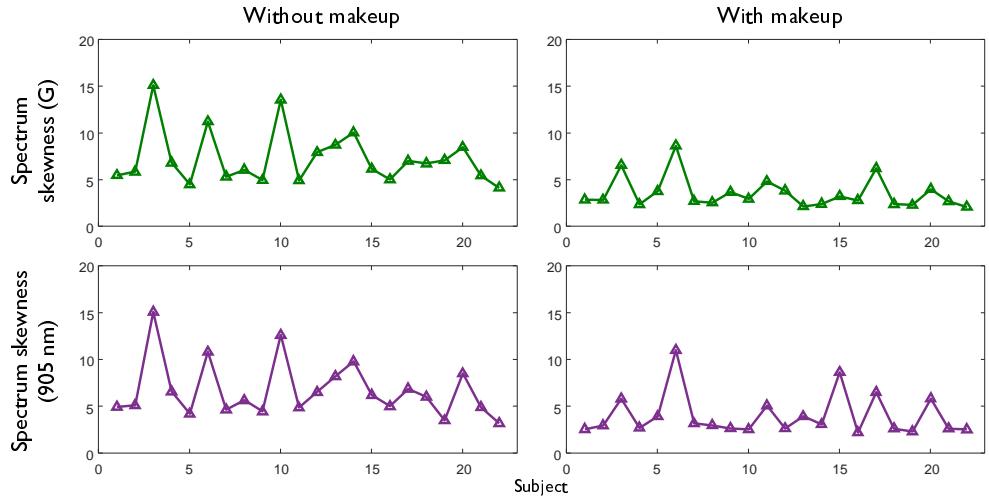


Fig. 9. The comparison of (a) signal spectrum of one subject (the subject exemplified in Fig. 8) and (b) spectrum skewness (i.e. the quality indicator) of 22 subjects, obtained in RGB (G channel) and NIR (905 nm), before and after applying the makeup.

less averaged skin pixels and thus higher quantization noise. It suggests that if makeup degrades the performance of remote-PPG extraction (for pulse-rate), (high-resolution) PPG imaging will suffer from more degradation. Therefore, remote-PPG based anti-spoofing [9–11] exploiting the living-skin detection/classification [31, 32] might be invalid in the use case of facial makeup (either in RGB or NIR), i.e. it cannot differentiate from a real face with makeup and a spoofing face (e.g. paper/replay attack). As anti-spoofing is typically applied in non-clinical scenarios like airport for security purpose, the challenge of makeup is, therefore, a non-negligible factor. Although emphasizing on non-clinical use cases, we also need to be careful/critical for the application in the clinical setting, i.e. we may boldly assume that the patients in the Intensive Care Unit (ICU) or Coronary Care Unit (CCU) have no makeup, but not the (suspected) patients in general ward or emergency department triage.

Additionally, we use a metric to evaluate the quality (e.g. noisiness) of raw wavelength signals.

Fig. 9 (a) exemplifies a comparison of the spectrum obtained on the subject of Fig. 8. It is clear that both the spectra of the G wavelength and 905 nm wavelength become noisier after applying the makeup. We also use the spectrum skewness as the quality metric [25] to evaluate the noisiness of the raw wavelength signals (see Fig. 9 (b)), i.e. higher spectrum skewness suggests better signal quality. It shows the raw signals in the G channel and 905 nm channel are indeed much noisier after applying the makeup. The trend of the spectrum skewness, obtained over 22 subjects, show that for the ones with stronger PPG signals (higher skewness without makeup), their signal quality (skewness) has less drops due to makeup.

Moreover, we notice that some subjects show acceptable pulse spectrograms after using makeup (in Appendix). More specifically, although the skin pulsatility of some subjects are reduced due to makeup, their pulsatile signatures still remain the correct channel order and thus do not completely fail the remote-PPG extraction algorithms. Our hypothesis for this observation is: (i) different subjects have different skin pulsabilities, which are related to their cardiovascular functions (e.g. blood pressure) and skin composition (e.g. melanin content), i.e. some have higher pulsatility than others; (ii) the actual amount of makeup used for each subject may vary a bit (though no major variation), and thus the influence of makeup is not exactly the same for all test subjects. These two factors may cause difference among these individuals. However, we mention that even though the average pulse rate (frequency information) remain after using makeup, correct blood oxygen saturation cannot be retrieved, as it requires very precise measurement of PPG amplitude from at least two wavelength channels.

Based on our earlier established research questions, we draw three conclusions correspondingly. Based on the makeup investigated in this study (type: skin protecting illuminating foundation Rose 01), we conclude that (1) it has negative influence on remote-PPG monitoring in both the RGB and NIR; (2) it has less influence in NIR than in RGB, which also validates our hypothesis, i.e. infrared light has deeper skin-penetration depth than visible light; (3) it makes the extraction of pulse rate more difficult and the extraction of blood oxygen saturation impossible. We consider there is no algorithmic solution (by image/signal processing) that can improve vital signs extraction from skin with makeup, as it is a physical limitation. A possible way to reduce erroneous measurement (or false alarms) is to first detect/recognize whether the subject has makeup on face. If makeup is being used, the system stops the measurement and/or suggests the subject to remove the makeup. At least, in this way, it will not confuse the system whether the measured abnormal vital signs (e.g. low skin pulsatility or low blood oxygen saturation) is due to human physiology (e.g. cardiac arrest, stroke) or makeup. Another option is to detect the skin region that does not apply makeup and select those areas to measure vital signs.

This paper focuses on studying the challenge of skin makeup for remote-PPG. There are other skin-related factors, such as facial hair, oil skin, sweat, etc., may also influence the measurement. These factors shall be studied in the future work.

4. Conclusion

In this paper, we perform a systematic study to investigate the impact of makeup on camera based PPG monitoring, in both the visible (RGB) and invisible (NIR) lighting conditions. We find that the makeup (investigated in this study, type: skin protecting illuminating foundation Rose 01) has negative influence on remote-PPG, which reduces the relative PPG strength (AC/DC) and changes the normalized PPG signature across multiple wavelengths. It makes (i) the pulse-rate extraction more difficult in both the RGB and NIR, though NIR is less affected than RGB; (ii) the blood oxygen saturation extraction in NIR impossible as the relative PPG contrast between different wavelengths (e.g. red and infrared) has been changed. Our study confirms that makeup is a threat for the application of camera based vital signs monitoring, which needs to be critically considered when customizing/deploying this technology, especially for non-clinical/non-patient scenarios where subjects may have makeup, such as driver monitoring in automotive, cardio-fitness training,

living-skin anti-spoofing, and video surveillance .

Acknowledgment

The authors would like to thank the volunteers from Philips Research and Eindhoven University of Technology for their participation in video recordings.

Funding

The Philips Research financially supported the research.

Disclosures

The authors declare that there are no conflicts of interest related to this article.

Appendix

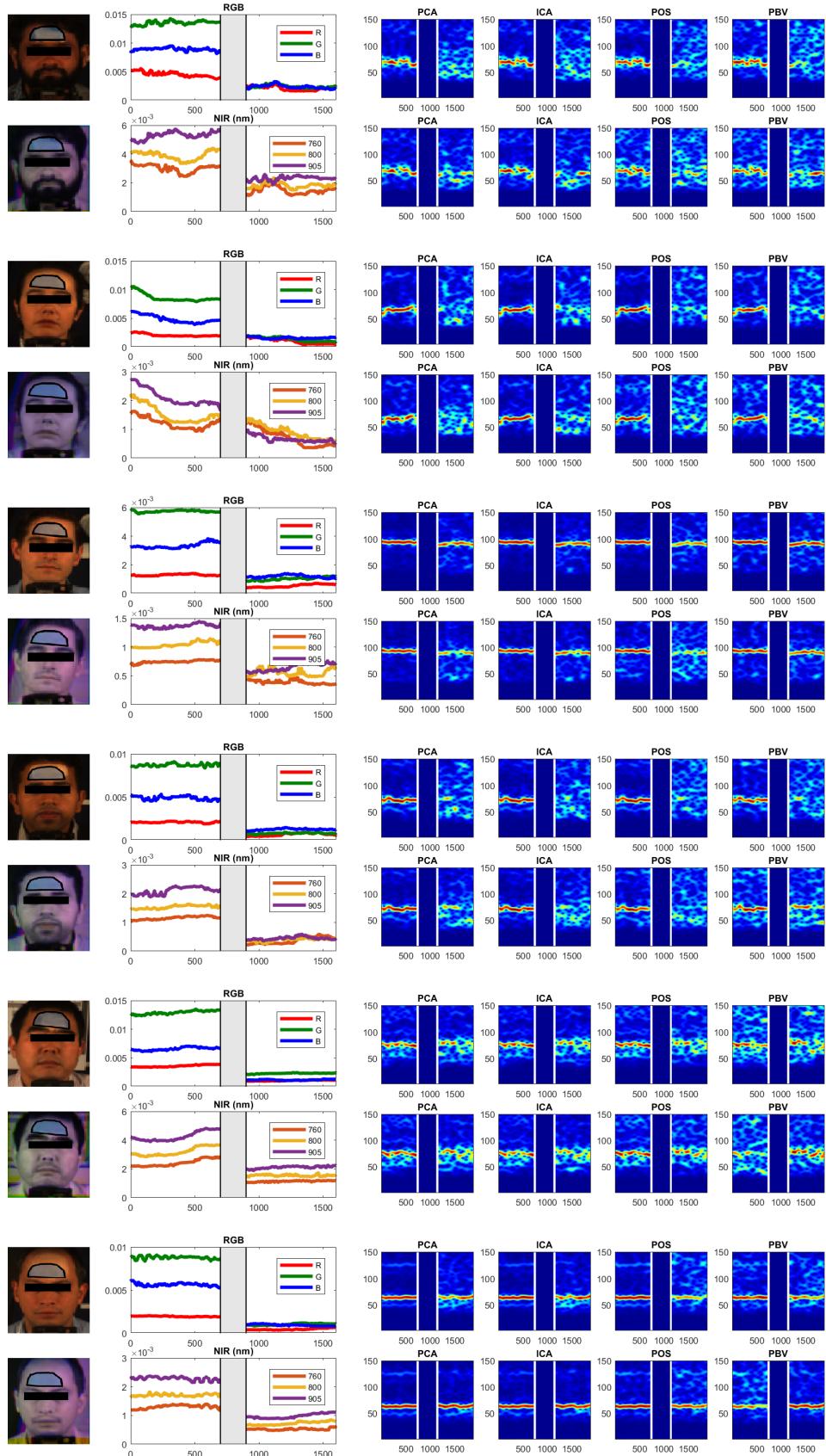


Fig. 10. The snapshot of recordings (RGB and NIR), relative PPG strength traces, and measured PPG spectrograms using four different remote-PPG algorithms.

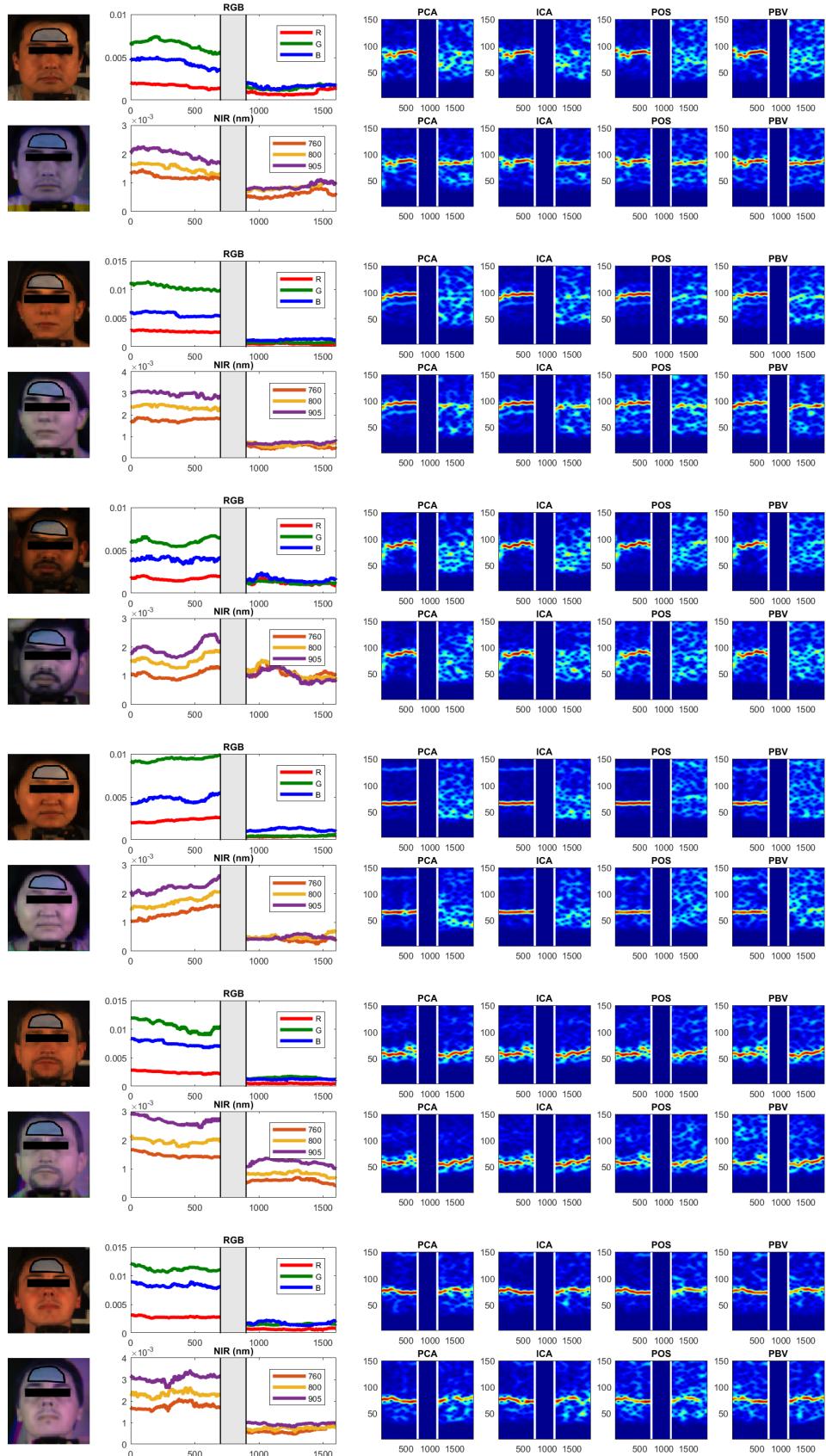


Fig. 11. The snapshot of recordings (RGB and NIR), relative PPG strength traces, and measured PPG spectrograms using four different remote-PPG algorithms.

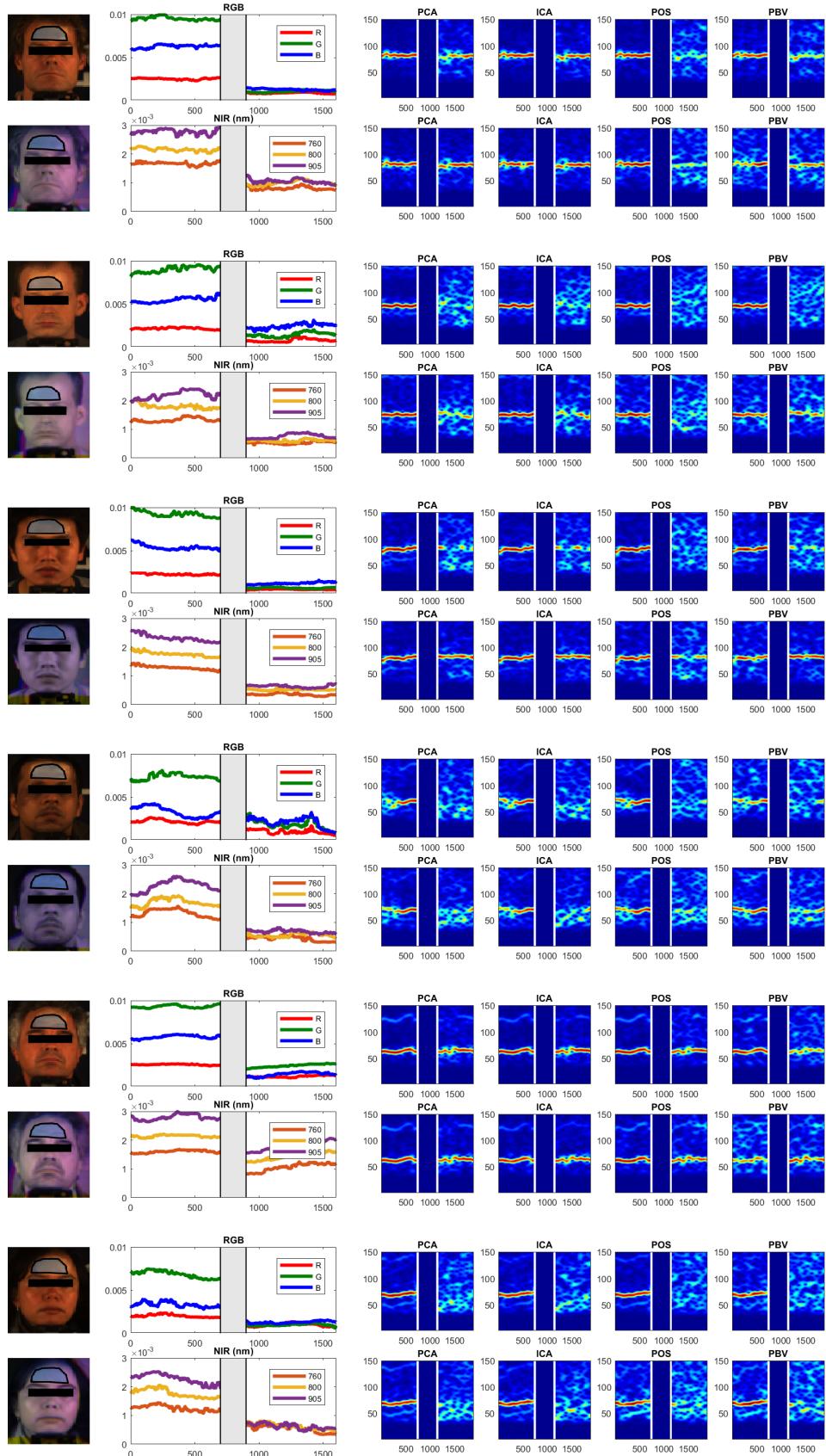


Fig. 12. The snapshot of recordings (RGB and NIR), relative PPG strength traces, and measured PPG spectrograms using four different remote-PPG algorithms.

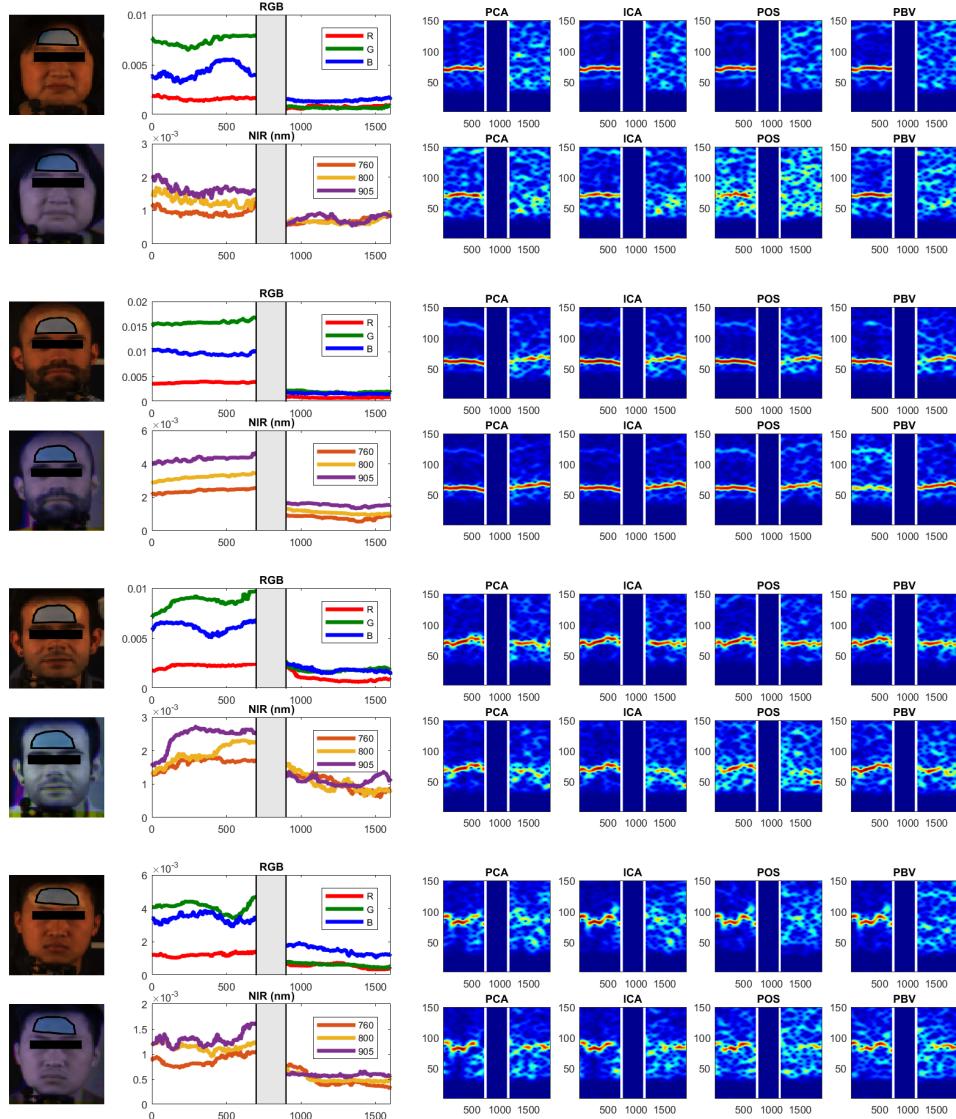


Fig. 13. The snapshot of recordings (RGB and NIR), relative PPG strength traces, and measured PPG spectrograms using four different remote-PPG algorithms.