

# Remote Heart Rate Measurement from RGB-NIR Video Based on Spatial and Spectral Face Patch Selection

Shiika Kado, *Student Member, IEEE*, Yusuke Monno, *Member, IEEE*, Kenta Moriwaki, Kazunori Yoshizaki, Masayuki Tanaka, *Member, IEEE*, and Masatoshi Okutomi, *Member, IEEE*

**Abstract**—In this paper, we propose a novel heart rate (HR) estimation method using simultaneously recorded RGB and near-infrared (NIR) face videos. The key idea of our method is to automatically select suitable face patches for HR estimation in both spatial and spectral domains. The spatial and spectral face patch selection enables us to robustly estimate HR under various situations, including scenes under which existing RGB camera-based methods fail to accurately estimate HR. For a challenging scene in low light and with light fluctuations, our method can successfully estimate HR for all 20 subjects ( $\pm 3$  beats per minute), while the RGB camera-based methods succeed only for 25% of the subjects.

## I. INTRODUCTION

Heart rate (HR) or pulse rate is one of the most essential vital signs, which provides the physiological and emotional state of a person. HR is typically measured using a photoplethysmographic (PPG) sensor attached to human skin. The optical PPG sensor measures light reflected from or transmitted through the skin. Since light intensity change on the skin measured over time is caused by blood volume change due to heartbeats, HR can be estimated from the PPG signal [1], [2].

It has been shown that HR can be estimated remotely using a digital camera [3], [4]. Similar to a PPG sensor, a camera can be used as a device to measure light intensity change on the skin. Therefore, HR can be estimated from a skin video typically on a face or a hand. Non-contact HR measurement from a video allows for various remote vital sensing applications such as monitoring of neonates [5] and prediction of daily health conditions [6].

Currently, most of camera-based HR estimation methods use a conventional RGB camera (see [2], [7], [8] for a review). However, RGB camera-based methods have a limitation that it is difficult to accurately estimate HR under low-light conditions or ambient light fluctuations. One way to overcome this limitation is to use an invisible near-infrared (NIR) light and an NIR camera. The feasibility of the NIR

camera-based method has been evaluated in several recent literatures [9], [10].

In this paper, to further improve robustness of camera-based HR estimation, we propose a novel HR estimation method using simultaneously recorded RGB and NIR face videos. In what follows, we refer to a set of spatially aligned RGB and NIR videos as an RGB-NIR video. The RGB-NIR video can be captured using a prism-based dual-CCD camera<sup>1</sup> or a single-sensor camera with an RGB-NIR filter array [11], [12]. Different from existing camera-based methods, our method using an RGB-NIR video covers both visible and NIR domains. This enables robust HR estimation under various illumination conditions, under which existing RGB or NIR camera-based methods fail to accurately estimate HR. Contributions of this paper are summarized as follows.

- We propose a novel HR estimation method using an RGB-NIR face video that is robust to various illumination conditions. To the best of our knowledge, our method is the first method that effectively combines RGB and NIR videos for remote HR measurement.
- We propose a novel HR estimation algorithm based on the automatic selection of suitable face regions (patches) both spatially and spectrally. Our algorithm enables accurate HR estimation without relying on heuristic selection of the face patches.
- We evaluate our method by experiments on 20 subjects and demonstrate that our method can estimate HR more robustly than existing methods.

## II. RELATED WORKS

While many approaches have been proposed for RGB camera-based HR estimation [13]–[27], we here focus on two approaches; (i) the approach using multiple channels within a single face region of interest (ROI) and (ii) the approach using a single channel, typically the G channel, within multiple face ROIs. Our method is built on the two approaches. For other approaches, we refer to comprehensive review papers [2], [7], [8].

As a representative method in the first approach, the Poh et al. method extracts temporal RGB intensity traces of the face, where each trace is calculated from averaged pixel intensities within a single face ROI [13], [14]. Then, this method employs independent component analysis (ICA) to extract the PPG signal from the RGB traces. Then, Fourier transform is applied to the extracted PPG signal to find the

This work was supported by the MIC/SCOPE #141203024.

S. Kado, Y. Monno, K. Moriwaki, M. Tanaka, and M. Okutomi are with the Department of Systems and Control Engineering, School of Engineering, Tokyo Institute of Technology, Meguro-ku, Tokyo 152-8550, Japan (e-mail: skado@ok.sc.e.titech.ac.jp; ymonno@ok.sc.e.titech.ac.jp; kmoriwaki@ok.sc.e.titech.ac.jp; mtanaka@sc.e.titech.ac.jp; mxo@sc.e.titech.ac.jp).

K. Moriwaki is now with Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan.

M. Tanaka is also with Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology, Koto-ku, Tokyo 135-0064, Japan.

K. Yoshizaki is with Olympus Corporation, Hachioji, Tokyo 192-8512, Japan (e-mail: kazunori.yoshizaki@ot.olympus.co.jp).

<sup>1</sup>AD-130GE, JAI Ltd., Japan. <http://www.jai.com/en/products/ad-130ge>.

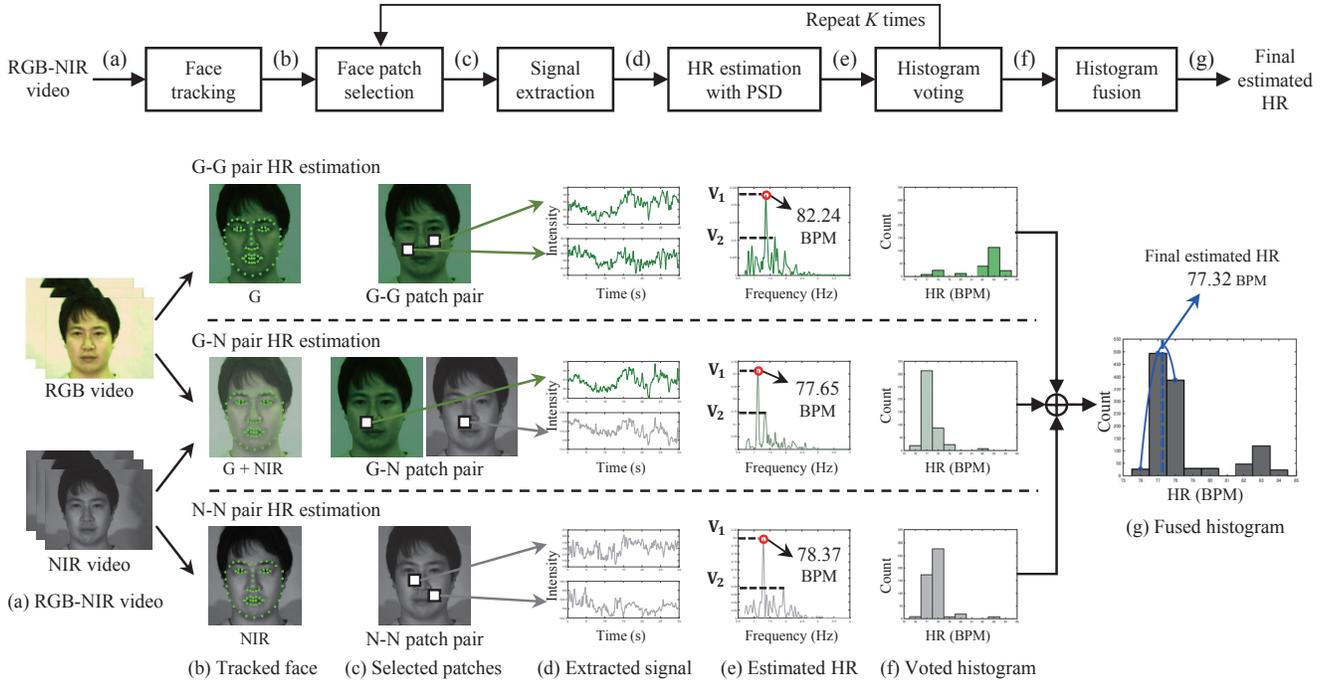


Fig. 1: Overall flow of our proposed HR estimation algorithm from an RGB-NIR face video.

most dominant frequency that is assumed to be the frequency of heartbeats. Some methods improved the Poh et al. method based on a better selection of the face ROI and channel combinations [15] or a machine learning approach [19].

As a method in the second approach, the Lam and Kuno’s method randomly selects two local face ROIs (patches) and extracts two temporal traces of the G channel intensity [23]. The two traces from the two face patches are used as inputs of ICA to extract the PPG signal. Similar to the Poh et al. method [13], [14], Fourier spectrum analysis of the extracted PPG signal is performed to estimate HR. To improve robustness to illumination variations at different face patches, random patch selection and HR estimation are repeatedly performed. The most reliable HR is finally selected based on a constructed histogram of the estimated HRs. Other methods have also been proposed to exploit spatial redundancy for selecting suitable face ROIs [24]–[26]. In these methods, the G channel is typically used because the G channel presents the strongest PPG amplitude [17], [28].

In our proposed method, we take a hybrid approach of the above two approaches. Specifically, we effectively use multiple (multispectral) channels and multiple face ROIs extracted from an input RGB-NIR face video. Our method is also different from existing non-RGB camera based methods using an NIR camera [9], [10], a thermal camera [29], or a multispectral camera [30]–[32], in that our method covers both visible and NIR domains and automatically selects suitable face ROIs both spatially and spectrally.

### III. PROPOSED HR ESTIMATION METHOD

#### A. Overview

Figure 1 shows the overall flow of our proposed HR estimation algorithm from an RGB-NIR face video. Our algorithm is inspired by the Lam and Kuno’s method [23] that randomly and repeatedly selects two face patches in the G channel (see G-G pair HR estimation in Fig. 1). The random and repeated patch selection effectively improves the robustness of HR estimation [23]. However, when using only the G channel, it is still difficult to accurately estimate HR under low-light conditions or ambient light fluctuations. This is the common limitation of the RGB camera-based methods.

To overcome the limitation, our algorithm also selects two face patches in the NIR channel (see N-N pair HR estimation in Fig. 1). To further improve the robustness and the versatility, our algorithm also combines the Poh et al. approach [13], [14] and uses both the G and the NIR channels<sup>2</sup> within a single face patch (see G-N pair HR estimation in Fig. 1). From those sets of spatially and spectrally selected patch pairs, our algorithm automatically selects suitable patch pairs for HR estimation based on the histogram fusion of reliably estimated HRs. Each processing step is detailed below.

#### B. Face landmark detection and tracking

Our algorithm first detects and tracks the face in the three videos, i.e., the G, the NIR, and the G+NIR videos. Here,

<sup>2</sup>Although our algorithm is generally extensible for any channel combinations, we here use only the G and the NIR channels because using the R and the B channels does not show improved performance in our experiments.

we assume that input RGB and NIR videos are spatially aligned. Thus, the G+NIR video is generated by the average of the G and the NIR videos. We use the algorithm in [33] (implemented in [34], [35]) to detect 66 face landmarks (see Fig. 1(b)) and the algorithm in [23] to track the face landmarks between image frames. If the landmarks cannot be detected in a frame, the last detected landmarks in previous frames are copied to the present frame.

### C. Face patch selection and signal extraction

Our algorithm then extracts three pairs of temporal intensity traces based on our proposed spatial and spectral face patch selection (see Fig. 1(c) and 1(d)). We refer to the three pairs as the G-G pair, the G-N pair, and the N-N pair, respectively. For each pair, temporal intensity traces are extracted as follows.

- 1) G-G pair: Two face patches are randomly selected in the G channel. For these patches, the temporal traces of the averaged G intensity are extracted.
- 2) G-N pair: One face patch is randomly selected in the G+NIR channel. For this patch, the temporal traces of the averaged G and NIR intensities are extracted.
- 3) N-N pair: Two face patches are randomly selected in the NIR channel. For these patches, the temporal traces of the averaged NIR intensity are extracted.

### D. HR estimation with power spectral density calculation

Our algorithm then follows the algorithm in [23] to estimate HR from each extracted pair of traces. First, moving average filter is applied to extracted intensity traces to reduce noise. Then, fast ICA [36] is performed to estimate the PPG signal from the intensity traces. Then, detrending filter [37] and moving average filter are applied to the estimated PPG signal to remove trends and reduce noise. Then, Welch's power spectral density (PSD) calculation [38] is applied to the PPG signal to find the most dominant frequency between 0.7Hz to 4Hz, which is assumed to be the frequency of the heartbeats (see Fig. 1(e)). The estimated HR is obtained in the form of beats per minute (BPM) by multiplying the corresponding frequency by 60.

### E. Histogram voting and fusion

To use only suitable face patch pairs for HR estimation, histogram voting is performed based on the reliability of the estimated HR [23]. The reliability is defined by the ratio of the spectral power of the most dominant frequency to that of the second dominant frequency, which is indicated as  $v_1/v_2$  in Fig. 1(e). This ratio implies how dominant the heartbeats frequency is. Thus, the higher ratio indicates more reliable HR estimation. When the ratio is higher than a threshold value  $T_r$ , the estimated HR is round to the integer value and voted to the corresponding histogram bin. The random patch selection and the histogram voting are repeated  $K$  times to construct the histogram of reliably estimated HRs.

As the result of the histogram voting, we obtain three histograms for the G-G, the G-N, and the N-N pairs, respectively (see Fig. 1(f)). To obtain the final estimated HR,

TABLE I: Illumination conditions for each scene.

	Fluorescent (FL) light	Illumination fluctuations	NIR light
Scene 1	600 lux	Without	ON
Scene 2	600 lux	Without	OFF
Scene 3	50 lux	Without	ON
Scene 4	50 lux	With	ON

histogram fusion is performed by adding the histograms whose maximum number of counts is more than a threshold value  $T_v$ . We regard the estimation as unreliable if the maximum number of counts is less than the threshold value. To estimate HR in the real value precision, parabola fitting is performed (see Fig. 1(g)). The peak of the fitted curve provides the final estimated HR.

## IV. EXPERIMENTAL RESULTS

We performed experiments under various illumination conditions to evaluate the performance of our method. Twenty subjects with both gender (three females) and different age (20's - 60's) took part in the experiments. The experiments were approved by the research ethics committees of Tokyo Institute of Technology and Olympus Corporation. The informed consent was obtained from all subjects.

We used a dual-CCD RGB-NIR camera (AD-130GE, JAI Ltd., Japan) to simultaneously record RGB and NIR face videos without misalignment. The face videos were recorded for 30 seconds with 1296×964 resolution, 12 bit depth and 30 frames per second. The subjects were asked to sit still in a chair, which was placed at a distance of 1.5 m from the camera. A contact PPG sensor (Procomp Infinity T7500M, Thought Technology Ltd., Canada) was attached to the subject's finger to obtain a reference HR.

To evaluate the robustness of our proposed method, face videos under four illumination conditions were captured. We used a fluorescent (FL) light for visible wavelengths and an NIR LED light for NIR wavelengths. We also used a display monitor, on which a movie with illumination fluctuations was played. The illumination conditions are listed in Table I and will be discussed in detail later.

We compared our method with the Poh et al. method [14] and the Lam et al. method [23]. The Poh et al. method uses a single face ROI and all RGB channels (noted as Poh-RGB). We also extended this method for the RGB-NIR video using all RGB and NIR channels (noted as Poh-RGBN). We used a manually selected cheek region as a fixed face ROI for the Poh-RGB and the Poh-RGBN methods. The Lam et al. method uses randomly selected two face ROIs and the G channel (noted as Lam-G), as explained before. We also applied this method to the NIR channel (noted as Lam-NIR) for comparison. Our proposed method performs the spatial and spectral face ROIs selection using the G and the NIR channels, as explained before. For the Lam-G, the Lam-NIR, and our methods, we used the same parameter values:

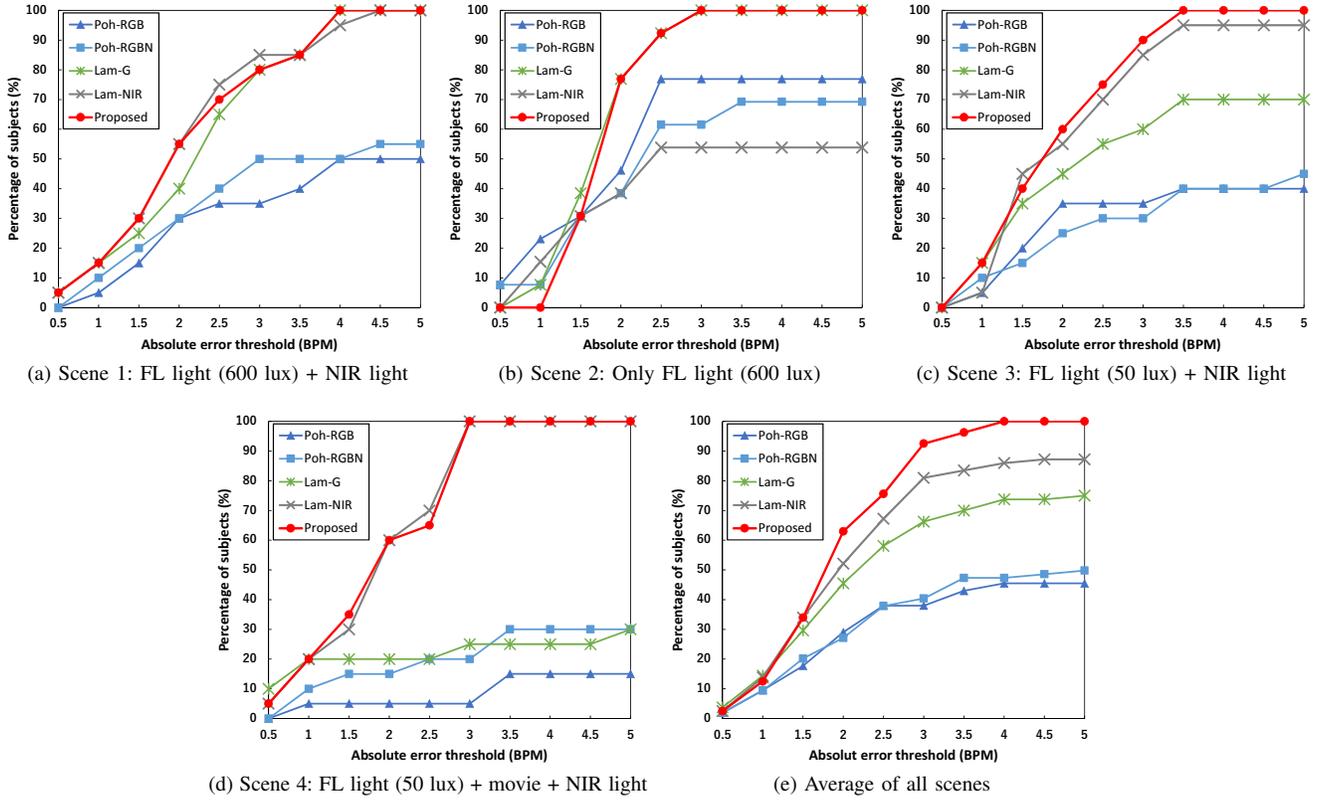


Fig. 2: Comparison of HR estimation accuracy.

$T_r = 2$ ,  $K = 500$ , and  $T_v = 10$ . The size of each face patch is set as 30% of the horizontal length of the detected face.

To focus on the differences of used face ROIs and used channels, for each method, we only changed the way of obtaining the intensity trace within the face ROI. All other processes, such as face tracking and the PPG signal estimation, are the same for all methods. If the face detection algorithm fails to detect face landmarks in more than 50% of all frames in a video, we did not perform HR estimation from that video. Failure cases also include the case that ICA does not converge.

Figure 2 shows the comparison of HR estimation accuracy under different illumination conditions. We evaluated the percentage of subjects that the absolute error in BPM is less than the threshold. Each result is discussed in detail below.

1) *Scene 1*: The videos were recorded under the FL light (600 lux) and the NIR light. Because this condition provides sufficient lighting for both the RGB and the NIR videos, the Lam-G, the Lam-NIR, and our methods provide similar results. These methods incorporate the random patch selection process, contributing to more accurate HR estimation than the Poh-RGB and the Poh-RGBN methods using the single fixed face ROI.

2) *Scene 2*: The videos were recorded under only the FL light (600 lux). This condition assumes the situation that the NIR light is not available or not necessary such as in daytime. In such the situation, our method automatically works as the RGB camera-based method, providing the similar result to

the result of the Lam-G method. In contrast, the Lam-NIR method does not work well for this scene because of the absence of the NIR light.

3) *Scene 3*: The videos were recorded under the FL light (50 lux) and the NIR light. This is a low-light condition, under which the RGB camera-based methods fail to accurately estimate HR. In contrast, the Lam-NIR method provides a better result by exploiting the NIR light. Our method also can successfully select suitable patch pairs from the G-N pair and the N-N pair and provide the best performance.

4) *Scene 4*: The videos were recorded under the FL light (50 lux) and the NIR light. To evaluate the robustness to illumination fluctuations, a movie had also been played on the display monitor, which was placed in front of a subject. This is a challenging condition for the RGB camera-based methods, because the videos were recorded under both a low-light condition and illumination fluctuations in the visible domain. Thus, it is difficult to accurately estimate HR by the Poh-RGB and the Lam-G methods. In contrast, our method can naturally work as the NIR camera-based method (similar to the Lam-NIR method) and provide the accurate result.

5) *Average of all scenes*: Figure 2(e) shows the average results of all scenes. The Lam-G and the Lam-NIR methods with the random patch selection are more robust than the Poh-RGB and the Poh-RGBN methods. However, the Lam-G and the Lam-NIR methods fail to accurately estimate HR at some scenes. In contrast, our method enables accurate HR estimation under various illumination conditions and

achieves the absolute errors less than 3 BPM for more than 90% of subjects.

## V. CONCLUSION

In this paper, we proposed a novel HR estimation method from an RGB-NIR face video based on the automatic spatial and spectral selection of suitable face patches. Experimental results for 20 subjects demonstrated that our method can robustly estimate HR under various illumination conditions, including the conditions under which existing RGB or NIR camera-based methods fail to accurately estimate HR. Since our method covers both visible and invisible NIR domains, a wide range of applications, such as day and night time HR monitoring, is expected in future.

## REFERENCES

- [1] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological Measurement*, vol. 28, no. 3, pp. R1–R39, 2007.
- [2] Y. Sun and N. Thakor, "Photoplethysmography revisited: From contact to noncontact, from point to imaging," *IEEE Trans. on Biomedical Engineering*, vol. 63, no. 3, pp. 463–477, 2016.
- [3] C. Takano and Y. Ohta, "Heart rate measurement based on a time-lapse image," *Medical Engineering & Physics*, vol. 29, no. 8, pp. 853–857, 2007.
- [4] W. Verkruyse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Optics Express*, vol. 16, no. 26, pp. 21 434–21 445, 2008.
- [5] L. A. Aarts, V. Jeanne, J. P. Cleary, C. Lieber, J. S. Nelson, S. B. Oetomo, and W. Verkruyse, "Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit - A pilot study," *Early Human Development*, vol. 89, no. 12, pp. 943–948, 2013.
- [6] J. Kranjec, S. Beguš, G. Geršek, and J. Drnovšek, "Non-contact heart rate and heart rate variability measurements: A review," *Biomedical Signal Processing and Control*, vol. 13, pp. 102–112, 2014.
- [7] A. Sikdar, S. K. Behera, and D. P. Dogra, "Computer-vision-guided human pulse rate estimation: A review," *IEEE Reviews in Biomedical Engineering*, vol. 9, pp. 91–105, 2016.
- [8] M. A. Hassan, A. S. Malik, D. Fofi, N. Saad, B. Karasfi, Y. S. Ali, and F. Mériaudeau, "Heart rate estimation using facial video: A review," *Biomedical Signal Processing and Control*, vol. 38, pp. 346–360, 2017.
- [9] M. van Gastel, S. Stuijk, and G. de Haan, "Motion robust remote-PPG in infrared," *IEEE Trans. on Biomedical Engineering*, vol. 62, no. 5, pp. 1425–1433, 2015.
- [10] R. Mitsuhashi, G. Okada, K. Kurita, K. Kagawa, S. Kawahito, C. Koopipat, and N. Tsumura, "Two-band infrared video-based measurement for non-contact pulse wave detection on face without visible lighting," *Proc. of Color and Imaging Conference (CIC)*, pp. 83–87, 2017.
- [11] Z. Chen, X. Wang, and R. Liang, "RGB-NIR multispectral camera," *Optics Express*, vol. 22, no. 5, pp. 4985–4994, 2014.
- [12] H. Teranaka, Y. Monno, M. Tanaka, and M. Okutomi, "Single-sensor RGB and NIR image acquisition: Toward optimal performance by taking account of CFA pattern, demosaicking, and color correction," *Proc. of IS&T Electronic Imaging*, pp. DPMI–256–1–6, 2016.
- [13] M. Z. Poh, D. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Optics Express*, vol. 18, no. 10, pp. 10 762–10 774, 2010.
- [14] —, "Advancements in non-contact, multiparameter physiological measurements using a webcam," *IEEE Trans. on Biomedical Engineering*, vol. 58, no. 1, pp. 7–11, 2011.
- [15] M. Lewandowska, J. Ruminski, T. Kocejko, and J. Nowak, "Measuring pulse rate with a webcam - A noncontact method for evaluating cardiac activity," *Proc. Federated Conf. on Computer Science and Information Systems (FedCSIS)*, pp. 405–410, 2011.
- [16] L. Wei, Y. Tian, Y. Wang, T. Ebrahimi, and T. Huang, "Automatic webcam-based human heart rate measurements using Laplacian eigenmap," *Proc. of Asian Conf. on Computer Vision (ACCV)*, pp. 281–292, 2012.
- [17] G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," *IEEE Trans. on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886, 2013.
- [18] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. Clifton, and C. Pugh, "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models," *Physiological Measurement*, vol. 35, no. 5, pp. 807–831, 2014.
- [19] H. Monkaresi, R. A. Calvo, and H. Yan, "A machine learning approach to improve contactless heart rate monitoring using a webcam," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1153–1160, 2014.
- [20] X. Li, J. Chen, G. Zhao, and M. Pietikäinen, "Remote heart rate measurement from face videos under realistic situations," *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 4321–4328, 2014.
- [21] G. de Haan and A. van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature," *Physiological Measurement*, vol. 35, no. 9, pp. 1913–1926, 2014.
- [22] L. Feng, L. M. Po, X. Xu, Y. Li, and R. Ma, "Motion-resistant remote imaging photoplethysmography based on the optical properties of skin," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 25, no. 5, pp. 879–891, 2015.
- [23] A. Lam and Y. Kuno, "Robust heart rate measurement from video using select random patches," *Proc. of IEEE Int. Conf. on Computer Vision (ICCV)*, pp. 3640–3648, 2015.
- [24] W. Wang, S. Stuijk, and G. de Haan, "Exploiting spatial redundancy of image sensor for motion robust rPPG," *IEEE Trans. on Biomedical Engineering*, vol. 62, no. 2, pp. 415–425, 2015.
- [25] 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.
- [26] S. Tulyakov, X. A. Pineda, E. Ricci, L. Yin, J. F. Cohn, and N. Seve, "Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions," *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2396–2404, 2016.
- [27] W. Wang, B. den Brinker, S. Stuijk, and G. de Haan, "Algorithmic principles of remote-PPG," *IEEE Trans. on Biomedical Engineering*, vol. 64, no. 7, pp. 1479–1491, 2016.
- [28] L. F. Corral, G. Paez, and M. Strojnik, "Optimal wavelength selection for non-contact reflection photoplethysmography," *Proc. of SPIE*, vol. 8011, pp. 801 191–1–7, 2011.
- [29] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE Trans. on Biomedical Engineering*, vol. 54, no. 8, pp. 1418–1426, 2007.
- [30] D. McDuff, S. Gontarek, and R. W. Picard, "Improvements in remote cardio-pulmonary measurement using a five band digital camera," *IEEE Trans. on Biomedical Engineering*, vol. 61, no. 10, pp. 2593–2601, 2014.
- [31] O. Gupta, D. McDuff, and R. Raskar, "Real-time physiological measurement and visualization using a synchronized multi-camera system," *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 46–53, 2016.
- [32] E. B. Blackford and J. R. Estep, "A multispectral testbed for cardiovascular sensing using imaging photoplethysmography," *Proc. of SPIE*, vol. 10072, pp. 100 720R–1–13, 2017.
- [33] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 1867–1874, 2014.
- [34] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.
- [35] Y. Nirkin, I. Masi, A. T. Trân, T. Hassner, and G. Medioni, "On face segmentation, face swapping, and face perception," 2017.
- [36] A. Hyvarinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, vol. 13, no. 4–5, pp. 411–430, 2000.
- [37] M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen, "An advanced detrending method with application to HRV analysis," *IEEE Trans. on Biomedical Engineering*, vol. 49, no. 2, pp. 172–175, 2002.
- [38] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.