ADAPTIVE FUSION OF RGB/NIR SIGNALS BASED ON FACE/BACKGROUND CROSS-SPECTRAL ANALYSIS FOR HEART RATE ESTIMATION

Kosuke Kurihara[1], *Daisuke Sugimura*[2] *and Takayuki Hamamoto*[1]

[1]Tokyo University of Science, Tokyo, 125-8585, Japan [2]Tsuda University, Tokyo, 187-8577, Japan

ABSTRACT

We propose a method for heart rate (HR) estimation that is robust to various situations such as bright, low-light, and varying illumination scenes. We capture temporal variations in the pixel values owing to person's cardiac pulse by using a camera that records red, green, and blue (RGB) and near-infrared (NIR) information. The key novelty of our method is to introduce a scheme for adaptive fusion of RGB and NIR signals for HR estimation, by analyzing variations in the background illuminations. RGB signals will be a good cue for HR estimation under bright scenes. In contrast, NIR signals are more reliable in HR estimation than RGB ones in complex illumination scenes, because NIR signals can be captured independent to changes in the background illuminations. By measuring correlations of signals between background and face regions, we adaptively utilize RGB and NIR signals for HR estimation. Experiments demonstrate the effectiveness of our method.

Index Terms— Remote vital sensing, RGB/NIR camera

1. INTRODUCTION

Heart rate (HR) is an important vital sign to estimate the physiological and the emotional states of person [1, 2]. Many methods for HR estimation using cameras have been proposed in the last decade [3–14]. Cameras enable to capture temporal changes in the skin color owing to person's cardiac pulse; thus, such time-series signals recorded in the videos can be exploited for HR estimation.

Previous video-based methods have shown their effectiveness in HR estimation. However, there remains a limitation to their use for uncontrolled scenes such as: low-light scenes, and scenes under varying illuminations (Fig.1). In low-light scenes, only small amount of light can come to camera. In scenes under varying illuminations, dominant color cast caused with background illuminations occurs on videos. It indicates that such uncontrolled illumination scenes make it difficult to measure changes in the skin color owing to cardiac pulse of person.

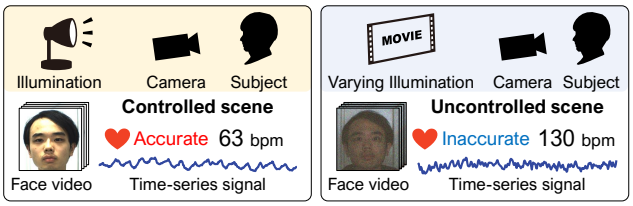


Fig. 1. Problems of HR estimation under uncontrolled scenes.

To overcome this problem, researchers have proposed methods for HR estimation using near-infrared (NIR) cameras [11–14]. By utilizing an NIR flash unit, sufficient amount of illuminations can be provided for scenes. Because NIR cameras can capture the NIR components only, influences of varying background illuminations can also be suppressed in HR estimation.

In fact, however, it is difficult to extract signals that correspond to HR from an NIR video, because of light absorption characteristics of blood [15, 16]. Hence, the accuracy of HR estimation using NIR signals would become lower than that of using RGB signals in bright scenes.

Kado *et al.* [14] used an RGB/NIR camera to estimate HRs under varying illuminations. They extended the existing method [7] by exploiting RGB and NIR patch signals measured in both spatial and spectral domains. They showed that the use of RGB and NIR signals enabled suppressions of influences of varying illuminations in HR estimation.

However, there still remains a problem for the method [14]. The authors of [14] simultaneously used RGB and NIR signals for HR estimation without considering background illuminations. When the background lights stably illuminate scenes, the use of NIR signals makes trouble for HR estimation because of light absorption characteristics of blood. In contrast, when background illumination fluctuates largely, the use of RGB signals decreases the accuracy of HR estimation.

In this study, we propose a novel method for remote HR sensing using RGB and NIR videos. The key novelty of the proposed method is to incorporate the measure of background illuminations into the HR estimation. When the background illumination fluctuates, it is preferable for HR estimation to

This article has been accepted for publication in IEEE International Conference on Image Processing (ICIP) 2019. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ICIP.2019.8803673

© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

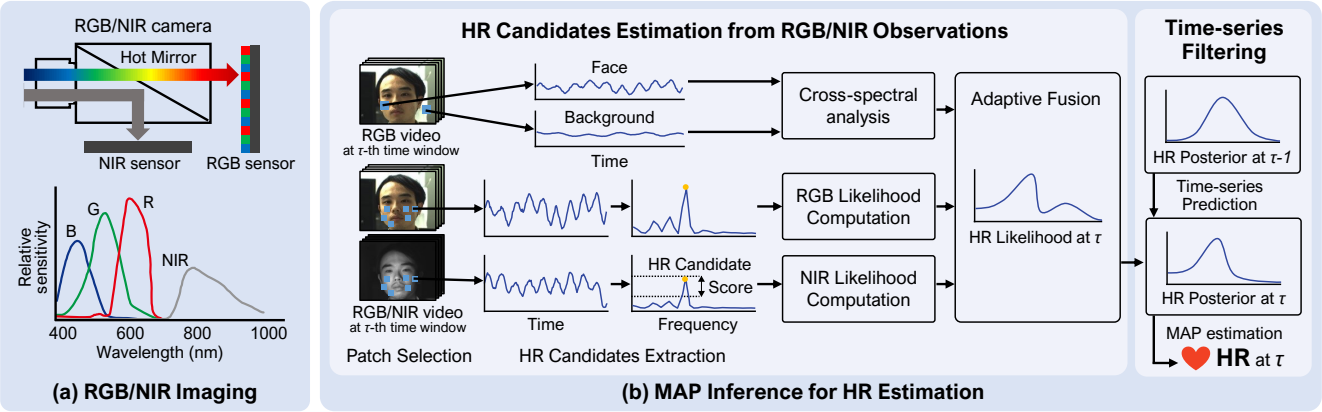


Fig. 2. Overview of our method: (a) RGB/NIR imaging; (b) HR inference with adaptive fusion of RGB/NIR signals.

emphasize contributions of NIR signals more than those of RGB ones. On the other hand, RGB signals will be more reliable for HR estimation than NIR in scenes where stable illuminations are maintained. By measuring correlations between face and background signals in the frequency domain, we can estimate which of RGB and NIR signals are more reliable for HR estimation. With this response, we can adaptively exploit time-series signals measured in RGB and NIR videos for HR estimation.

2. OVERVIEW OF THE PROPOSED METHOD

Our goal is to achieve accurate HR estimation under varying illuminations. Figure 2 shows an overview of the proposed method. We perform a time-series filtering to build a prior probability of HR. We then compute a posterior probability of HR by calculating a likelihood using the RGB and NIR observations. In this processing, we adaptively fuse RGB and NIR observations via cross-spectral analysis of face and background signals. We finally infer the latent HR based on a maximum a posteriori (MAP) framework.

2.1. Problem Formulation

Let the latent HR at the τ-th time window be h_τ, and the observed pair of RGB and NIR videos at the τ-th time window be $\mathbf{z}_\tau = \left(\mathbf{I}_\tau^{\mathrm{RGB}}, \mathbf{I}_\tau^{\mathrm{NIR}}\right)$, where the RGB and NIR videos are represented as $\mathbf{I}_\tau^{\mathrm{RGB}}$ and $\mathbf{I}_\tau^{\mathrm{NIR}}$, respectively. The size of time window is denoted as N. We also define all of the observations up to the τ-th time window as $\mathcal{Z}_{1:\tau} = (\mathbf{z}_1, ..., \mathbf{z}_\tau)$.

The posterior probability of h_τ, defined as $p(h_\tau \mid \mathcal{Z}_{1:\tau})$, can be derived using Bayes rule as

$$\begin{aligned} p(h_\tau \mid \mathcal{Z}_{1:\tau}) &\propto p(\mathbf{z}_\tau \mid h_\tau)\, p(h_\tau \mid \mathcal{Z}_{1:\tau-1}) \\ &= p(\mathbf{z}_\tau \mid h_\tau) \int p(h_\tau \mid h_{\tau-1})\, p(h_{\tau-1} \mid \mathcal{Z}_{1:\tau-1})\, \mathrm{d}h_{\tau-1}\,, \end{aligned} \quad (1)$$

where $p(\mathbf{z}_\tau \mid h_\tau)$, $p(h_\tau \mid h_{\tau-1})$, and $p(h_{\tau-1} \mid \mathcal{Z}_{1:\tau-1})$ represent a likelihood at τ, a state transition probability from $\tau-1$ to τ, and a posterior probability at $\tau-1$, respectively. In our Bayesian inference, we utilize a particle filter framework [17] to estimate the posterior probability $p(h_\tau \mid \mathcal{Z}_{1:\tau})$. In Eq. (1), we model the state transition probability $p(h_\tau \mid h_{\tau-1})$ as a Gaussian distribution (i.e., the first-order autoregressive (AR) model) with the mean $\mu\ (= h_\tau)$ and the standard deviation σ. The likelihood $p(\mathbf{z}_\tau \mid h_\tau)$ is computed using the RGB and NIR video observations. We describe the details of how to compute $p(\mathbf{z}_\tau \mid h_\tau)$ in the subsequent sections.

We infer the latent HR h_τ^* based on MAP estimation framework; it is obtained as

$$h_\tau^* = \arg\max_{h_\tau}\ p(h_\tau \mid \mathcal{Z}_{1:\tau})\,. \quad (2)$$

3. ADAPTIVE FUSION OF RGB AND NIR SIGNALS

We model the likelihood term in Eq. (1) using the RGB and NIR observations. Specifically, it is represented as

$$\begin{aligned} p(\mathbf{z}_\tau \mid h_\tau) = \gamma_\tau \beta_\tau^{\mathrm{back}}\, p_{\mathrm{RGB}}(\mathbf{I}_\tau^{\mathrm{RGB}} \mid h_\tau) \\ +(1-\gamma_\tau \beta_\tau^{\mathrm{back}})\, p_{\mathrm{NIR}}(\mathbf{I}_\tau^{\mathrm{NIR}} \mid h_\tau)\,, \end{aligned} \quad (3)$$

where $p_{\mathrm{RGB}}(\mathbf{I}_\tau^{\mathrm{RGB}} \mid h_\tau)$ represents a likelihood that is computed exploiting an RGB video, whereas $p_{\mathrm{NIR}}(\mathbf{I}_\tau^{\mathrm{NIR}} \mid h_\tau)$ denotes the one that is computed exploiting an NIR video. The weight γ_τ assesses which signals measured from the RGB and NIR videos are more reliable. The other weight $\beta_\tau^{\mathrm{back}}$ is computed based on cross-correlations of signals in the face and background regions.

3.1. Face Patch Selection

According to past studies [7, 8], it is reported that analyzing time-series signals in local image patch basis was effective for HR estimation. Based on this findings, we utilize local image patches for HR estimation.

We first perform a face detection [18, 19] to extract cheek regions in the subject's face. This is primarily because local patches selected from a cheek region allow the extraction of reliable signals for HR estimation, as reported in [7, 10, 14].

We then randomly select local patches from the cheek region. We denote a set of the selected patch sequences in an

RGB video as $\mathcal{P}_\tau^{\mathrm{RGB}} = \{\mathbf{p}_{\tau,i}^{\mathrm{RGB}}\}_{i=1}^{M}$ (M is the number of local patches extracted), where $\mathbf{p}_{\tau,i}^{\mathrm{RGB}}$ is an $n \times n \times N$ -sized patch image sequence in the τ-th time window. We obtain $\mathbf{p}_{\tau,i}^{\mathrm{RGB}}$ by estimating visual correspondence between consecutive frames using facial landmark positions.

We also randomly extract such local patch sequences from the NIR video in the same manner. We define them as $\mathcal{P}_\tau^{\mathrm{NIR}} = \{\mathbf{p}_{\tau,i}^{\mathrm{NIR}}\}_{i=1}^{M}$.

3.2. HR Candidates Extraction from RGB Video

3.2.1. Color Difference Space Conversion

According to the past study [9], it was shown that analysis in the color difference space was effective for HR estimation from an RGB video. Based on this findings, we project the RGB values in the local patches into the color difference space, and then measure the temporal variations in each patch.

3.2.2. Frequency Analysis

We perform a fast Fourier transform (FFT) to obtain a power spectrum of the time-series patch signal. Furthermore, we utilize a band-pass filter that filters out the spectrum component in the frequency range 0.7 - 2.5 Hz. We determine these cut-off frequencies based on the knowledge about the range of HR of normal person [20]. We apply this band-pass filter to a power spectrum measured from a local patch. We define a power spectra filtered by this band-pass filter for the i-th patch as $\mathbf{B}_{\tau,i}^{\mathrm{RGB}}$.

3.2.3. HR Candidates Extraction

We explore the frequency that has the largest power spectrum component (defined as v_1) in $\mathbf{B}_{\tau,i}^{\mathrm{RGB}}$, which is assumed to be the frequency of cardiac pulse of person [7, 14]. We denote such frequency as $b_{\tau,i}^{\mathrm{RGB}*}$. As in [7, 14], we consider a confidence score that assesses how much the HR candidate $b_{\tau,i}^{\mathrm{RGB}*}$ is reliable. To do this, we explore the second-largest power spectrum component v_2 from $\mathbf{B}_{\tau,i}^{\mathrm{RGB}}$. We then compute the confidence score as $\alpha_{\tau,i}^{\mathrm{RGB}} = v_1/v_2$.

We perform the above processing to all of the local patches. We finally obtain a pair of the HR candidates and the corresponding confidence scores for all of the selected patches. We respectively denote them as $(b_{\tau,1}^{\mathrm{RGB}*}, ..., b_{\tau,M}^{\mathrm{RGB}*})$, and $(\alpha_{\tau,1}^{\mathrm{RGB}}, ..., \alpha_{\tau,M}^{\mathrm{RGB}})$.

3.3. HR Candidates Extraction from NIR Video

Similar to the procedure described in the former section, we extract candidates of HR from the NIR video. We apply the frequency analysis (Sect. 3.2.2) to each local NIR patch in $\mathcal{P}_\tau^{\mathrm{NIR}}$. Following the procedure described in Sect. 3.2.3, we then extract the HR candidates from $\mathcal{P}_\tau^{\mathrm{NIR}}$ that are defined as $(b_{\tau,1}^{\mathrm{NIR}*}, ..., b_{\tau,M}^{\mathrm{NIR}*})$. The corresponding confidence scores are represented as $(\alpha_{\tau,1}^{\mathrm{NIR}}, ..., \alpha_{\tau,M}^{\mathrm{NIR}})$.

3.4. Likelihood Computation

Using $\{b_{\tau,i}^{\mathrm{RGB}*}\}$, $\{\alpha_{\tau,i}^{\mathrm{RGB}}\}$ and $\{b_{\tau,i}^{\mathrm{NIR}*}\}$, $\{\alpha_{\tau,i}^{\mathrm{NIR}}\}$, we respectively model the RGB and NIR likelihoods $p_{\mathrm{RGB}}(\mathbf{I}_\tau^{\mathrm{RGB}} \mid h_\tau)$ and $p_{\mathrm{NIR}}(\mathbf{I}_\tau^{\mathrm{NIR}} \mid h_\tau)$ based on a weighted kernel density estimation. We compute $p_D(\mathbf{I}_\tau^D \mid h_\tau)$ ($D \in \{\mathrm{RGB}, \mathrm{NIR}\}$) as

$$p_D(\mathbf{I}_\tau^D \mid h_\tau) = \frac{1}{\sum_{i=1}^{M} \alpha_{\tau,i}^D} \frac{1}{MW} \sum_{i=1}^{M} \alpha_{\tau,i}^D K\left(\frac{h_\tau - b_{\tau,i}^{D}*}{W}\right), \quad (4)$$

where $K(\cdot)$ represents a Gaussian kernel with the bandwidth W. Using the sampling rate of camera f_s, we set W such that $W = f_s/N$ (i.e., frequency resolution).

3.5. Modeling Influences of Background Illuminations

To examine influences of background illuminations in estimating HR, we compute the weight $\beta_\tau^{\mathrm{back}}$ by measuring correlations of signals of background and those of face of person.

3.5.1. Cross-spectral Analysis

We randomly select local image patches from face and background regions in the RGB video, and extract temporal variations in these patches. In every color channel, we average the extracted time-series signals over the face patches, and then remove its trend component from the averaged patch signal (i.e., subtraction of temporal mean value). This processing allows us to analyze similarity between harmonic components of signals. In the same manner, We average and detrend patch signals extracted from the background regions.

In order to measure the spectral correlations between the face and background signals, we apply a cross-spectral analysis to the obtained face and background signals. We denote the obtained cross-power spectra as $\{\mathbf{S}_\tau^c\}_{c\in\mathrm{R,G,B}}$, where each component $\mathbf{S}_\tau^c$ has N frequency components: $\mathbf{S}_\tau^c = (S_{\tau,f_1}^c, ..., S_{\tau,f_N}^c)$.

3.5.2. Computation of Weight of Background Illuminations

Using $\{\mathbf{S}_\tau^c\}$, we compute $\beta_\tau^{\mathrm{back}}$. When the cross-power spectra have strong components in the range of HR relative to the other frequency components, we consider that differentiations between HR and background signals will be difficult, because they are highly correlated with each other.

We assess how much the background illuminations will influence on the HR estimation. For each color channel, we first compute the relative largest cross-power spectrum m^c in the range of possible HR against the maximum cross-power spectrum ; it is computed as

$$m^c = \max_{f \in R_{\mathrm{HR}}} \frac{S_{\tau,f}^c}{\max_f S_{\tau,f}^c} \quad c \in \{\mathrm{R}, \mathrm{G}, \mathrm{B}\}, \quad (5)$$

where R_{HR} is a range of possible HR (from 0.7 to 2.5 Hz).

We compute $\beta_\tau^{\mathrm{back}}$ by utilizing the maximum response among $m^{\mathrm{R}}, m^{\mathrm{G}}$ and m^{B} as

$$\beta_\tau^{\mathrm{back}} = 1 - \max_{c\in\{\mathrm{R,G,B}\}} m^c. \quad (6)$$

3.6. Cross-domain Reliability Computation

We assess which of RGB and NIR face patch signals will be more reliable in HR estimation. When the RGB confidence scores $\{\alpha_{\tau,i}^{\mathrm{RGB}}\}$ show higher values than the NIR ones $\{\alpha_{\tau,i}^{\mathrm{NIR}}\}$, we consider that RGB patch signals are more reliable than NIR. Therefore we compute the score γ_τ that characterizes such reliability as

$$\gamma_\tau = \frac{\sum_{i=1}^{M} \alpha_{\tau,i}^{\mathrm{RGB}}}{\sum_{i=1}^{M} (\alpha_{\tau,i}^{\mathrm{RGB}} + \alpha_{\tau,i}^{\mathrm{NIR}})} . \quad (7)$$

4. EXPERIMENTS

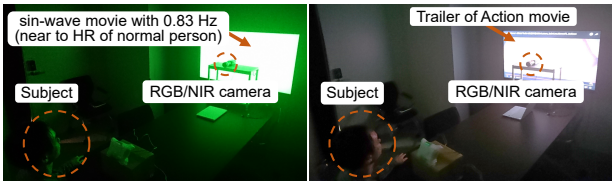
To evaluate the effectiveness of our method, we conducted experiments using real RGB and NIR videos. 11 subjects (22-24 years old) participated, and they were asked to sit still.

We tested our method under the following illumination conditions: (i) bright scene (illumination level: 600 lx), (ii) low-light scene (illumination level: 0.4 lx), (iii) scene under varying illumination with a frequency near to HR of normal person (illumination level: 1 lx), and (iv) realistic scene like a theater (low-light scene with varying illuminations; illumination level: 1 lx). We refer to these conditions to "S1", "S2", "S3" and "S4", respectively. In "S3" and "S4", we respectively played a sine-wave movie with 0.83 Hz (50 bpm), and a YouTube movie [1] (Trailer of Action movie) in a display located near to subjects (Fig.3). We captured RGB raw and NIR videos for 2 minutes with a two-plate RGB/NIR camera (JAI AD-130GE). The 8 bit RGB raw and NIR images with 1296×966 resolutions were recorded at 30 fps. For each RGB raw frame, we applied a demosaicing method [21] to obtain an RGB image with full resolutions. We used a pulse oximeter (CONTEC CMS50D+) to obtain a reference HR.

We compared our method with state-of-the-art methods: Lam *et al.* [7] (method using an RGB video), method [7] using an NIR video, and Kado *et al.* [14] (method using RGB and NIR videos). By conducting preliminary experiments using the other RGB and NIR videos, we set parameters for our method such that $\sigma^2 = 25$, $N = 900$ (30 sec), $n = 20$, and $M = 160$. We used the same parameters for all of the evaluations. For the comparison methods, we used the same image patches, and set their parameters such that they enabled to obtain the best results. We quantitatively evaluated the results by using both mean absolute error (MAE) and root mean squared error (RMSE) measures.

Table 1 lists the comparison results. Each value was obtained by averaging results for all of the subjects. We can see that our method outperformed the other methods in all of the scenes. Figure 4 shows the Bland-Altman plot [22,23], which is a method of data plotting used for assessing agreement between results obtained with two measurement methods; plots where the measurements are distributed narrowly around zero

[1] https://www.youtube.com/watch?v=zJO4LPkcadM



(a) scene "S3"(1 lx) (b) scene "S4"(1 lx)

Fig. 3. Illumination conditions ("S3" and "S4").

Table 1. Quantitative comparisons using the averaged MAE and RMSE measures. The best scores are represented in **bold**.

	MAE				RMSE			
	S1	S2	S3	S4	S1	S2	S3	S4
[7]	3.9	14.4	22.7	25.7	5.9	17.9	22.8	26.4
[7] (NIR)	6.7	6.3	4.2	6.8	10.1	10.8	7.1	8.6
[14]	4.9	7.0	22.6	14.1	7.4	11.4	22.7	18.5
Ours	**1.9**	**2.7**	**2.5**	**2.6**	**2.4**	**3.9**	**3.3**	**3.5**

indicate the better performance. In this assessment, we compared the measurements obtained with our method and those with a pulse oximeter (i.e., reference HR). We can see that our method can estimate more accurately than the other methods.

We finally show temporal changes in the estimated HR for subject #6 in scene "S3" in Fig.5. We can see that temporally-stable and accurate results were obtained using our method.

5. CONCLUSION

We proposed a method for HR estimation that is robust to various illumination conditions. We utilized an RGB/NIR camera to capture temporal variations in the pixel value owing to person's cardiac pulse. By analyzing correlations of signals between background and face regions, we adaptively utilized the RGB and NIR signals for HR estimation. Through the experiments, we demonstrated the effectiveness of our method.

6. REFERENCES

[1] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, 2007.

[2] G. Valenza, L. Citi, A. Lanatrk, E. P. Scilingo, and R. Barbieri, "Revealing real-time emotional responses: a personalized assessment based on heartbeat dynamics," *Sci. Rep.*, vol. 4, 2014.

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

[4] 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," *Physiol. Meas.*, vol. 35, no. 5, pp. 807, 2014.

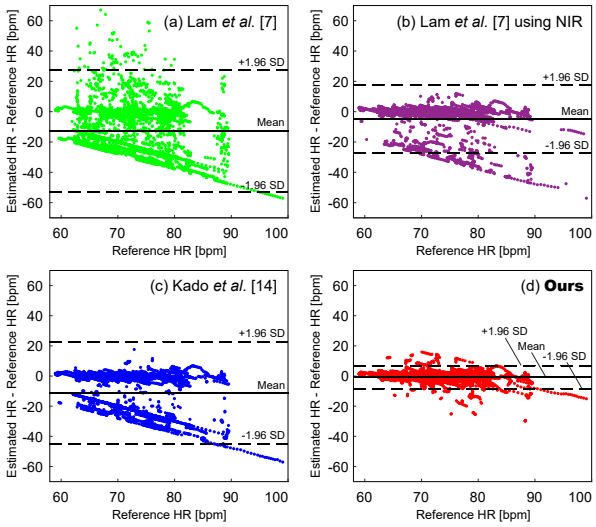


Fig. 4. Quantitative Comparisons using Bland-Altman plots for all of the scenes and subjects. In each figure, solid line shows the mean error, and dashed lines indicate 95% limits of agreement between the estimated and reference HR.

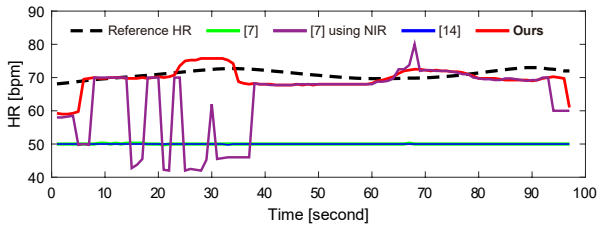


Fig. 5. Temporal changes in the estimated HRs for subject #6 in scene "S3".

[5] F. Zhao, M. Li, Y. Qian, and J. Z. Tsien, "Remote measurements of heart and respiration rates for telemedicine," *PLoS One*, vol. 8, no. 10, pp. 1–14, 2013.

[6] W. Wang, A. C. D. Brinker, and G. D. Haan, "Single element remote-ppg," *IEEE TBE*, 2018, Early Access.

[7] A. Lam and Y. Kuno, "Robust heart rate measurement from video using select random patches," in *Proc. of IEEE ICCV*, Dec. 2015, pp. 3640–3648.

[8] M. Kumar, A. Veeraraghavan, and A. Sabharwal, "DistancePPG: Robust non-contact vital signs monitoring using a camera," *Biomed. Opt. Express*, vol. 6, no. 5, pp. 1565–1588, 2015.

[9] W. Wang, A. C. D. Brinker, S. Stuijk, and G. D. Haan, "Algorithmic principles of remote ppg," *IEEE TBE*, vol. 64, no. 7, pp. 1479–1491, 2017.

[10] L. Feng, L. Po, X. Xu, Y. Li, and R. Ma, "Motion-resistant remote imaging photoplethysmography based on the optical properties of skin," *IEEE TCSVT*, vol. 25, no. 5, pp. 879–891, 2015.

[11] M. V. Gastel, S. Stuijk, and G. D. Haan, "Motion robust remote-ppg in infrared," *IEEE TBE*, vol. 62, no. 5, pp. 1425–1433, 2015.

[12] E. M. Nowara, T. K. Marks, H. Mansour, and A. Veeraraghavan, "Sparseppg: Towards driver monitoring using camera-based vital signs estimation in near-infrared," in *Proc. of IEEE CVPRW*, June 2018, pp. 1272–1281.

[13] S. B. Park, G. Kim, H. J. Baek, J. H. Han, and J. H. Kim, "Remote pulse rate measurement from near-infrared videos," *IEEE SPL*, vol. 25, no. 8, pp. 1271–1275, 2018.

[14] S. Kado, Y. Monno, K. Moriwaki, K. Yoshizaki, M. Tanaka, and M. Okutomi, "Remote heart rate measurement from rgb-nir video based on spatial and spectral face patch selection," in *Proc. of IEEE EMBC*, July 2018, pp. 5676–5680.

[15] L. F. C. Martinez, G. Paez, and M. Strojnik, "Optimal wavelength selection for noncontact reflection photoplethysmography," in *Proc. of SPIE ICO*, Nov. 2011, vol. 8011, pp. 801191–1–801191–7.

[16] E. B. Blackford, J. R. Estepp, and D. McDuff, "Remote spectral measurements of the blood volume pulse with applications for imaging photoplethysmography," in *Proc. of SPIE BiOS*, Feb. 2018, vol. 10501, pp. 105010Z–1–105010Z–8.

[17] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE TSP*, vol. 50, no. 2, pp. 174–188, 2002.

[18] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. of IEEE CVPR*, June 2005, pp. 886–893.

[19] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc. of IEEE CVPR*, June 2014, pp. 1867–1874.

[20] P. Palatini, "Need for a revision of the normal limits of resting heart rate," *Hypertension*, vol. 33, no. 2, pp. 622–625, 1999.

[21] H. S. Malvar, L. He, and R. Cutler, "High-quality linear interpolation for demosaicing of bayer-patterned color images," in *Proc. of IEEE ICASSP*, May 2004, vol. 3, pp. 485–488.

[22] J. M. Bland and D.Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, no. 8476, pp. 307 – 310, 1986.

[23] J. S. Krouwer, "Why bland-altman plots should use x, not (y+x)/2 when x is a reference method," *Stat. Med.*, vol. 27, no. 5, pp. 778–780, 2008.