RealSense = Real Heart Rate: Illumination Invariant Heart Rate Estimation from Videos

Jie Chen[1], Zhuoqing Chang[2], Qiang Qiu[2], Xiaobai Li[1], Guillermo Sapiro[2], Alex Bronstein[3], Matti Pietikäinen[1]
[1]University of Oulu, Finland
[2]Duke University, USA
[3]Tel Aviv University, Israel

***Abstract*—Recent studies validated the feasibility of estimating heart rate from human faces in RGB video. However, test subjects are often recorded under controlled conditions, as illumination variations significantly affect the RGB-based heart rate estimation accuracy. Intel newly-announced low-cost RealSense 3D (RGBD) camera is becoming ubiquitous in laptops and mobile devices starting this year, opening the door to new and more robust computer vision. RealSense cameras produce RGB images with extra depth information inferred from a latent near-infrared (NIR) channel. In this paper, we experimentally demonstrate, for the first time, that heart rate can be reliably estimated from RealSense near-infrared images. This enables illumination invariant heart rate estimation, extending the heart rate from video feasibility to low-light applications, such as night driving. With the (coming) ubiquitous presence of RealSense devices, the proposed method not only utilizes its near-infrared channel, designed originally to be hidden from consumers; but also exploits the associated depth information for improved robustness to head pose.**

***Keywords*—Heart Rate Estimation; Illumination Invariant; RealSense**

I. INTRODUCTION

Heart rate (pulse) is a measure of the number of heart beats in a minute. It is a critical vital sign to assess the physiological state of a subject. In many applications, it is preferred or even required to measure the heart rate, e.g., of a patient or a driver, in a passive and remote manner. Hear rate (HR) monitoring has many potential applications. With the ability to 'see' inner changes like the heartbeat, video processing research can be broadened in many ways. Specifically, it can be used for real-time remote medical examinations and support long-term HR monitoring. It can also be used for affective computing. In addition to the analysis of explicit behaviours like poses and gestures, inner physiological changes provide information for better understanding people's behaviour.

Traditional HR measurement methods are divided into two categories. The first one relies on special electronic or optical sensors, and most of the instruments require skin-contact, making them inconvenient and uncomfortable, e.g., Electrocardiography (ECG). The second one is to use photoplethysmography (PPG) [5, 10, 11]. The principle of PPG method is to illuminate the skin with a light-emitting diode (LED) and then measure the amount of light reflected or transmitted to a photodiode. Although it is possible to use PPG based settings to measure HR without any contact, this method still requires special lighting sources and sensors.

Recent remote HR monitoring methods include motion-based method [2, 26] and color-based methods using ordinary commercial cameras [12, 14, 15, 16, 23, 28, 30]. For the motion-based method, the cyclical movement of blood from the heart to the head via the abdominal aorta and the carotid arteries causes the head to move or face color to vary in a periodic way. A typical motion-based method was proposed by Balakrishnan et al [2]. They tracked subtle head oscillations caused by cardiovascular circulation, and used principal component analysis (PCA) to extract the pulse signal from the trajectories of multiple tracked feature points. The motion-based method requires the recorded subjects to be strictly stationary and sit upright for the duration of the video, which is often difficult to enforce in real scenarios.

For the color-based methods, Poh et al. explored the possibility to measure HR from face videos recorded by a web-cam [15]. They detected the region of interest (ROI, i.e. the face area) and computed the mean pixel values of the ROI of each frame from three color channels for HR measurement. Later they improved their method by adding several temporal filters both before and after applying independent component analysis (ICA) [16]. The advanced ICA method achieved very high accuracy for measuring HR on their data. Kwon et al. recorded face videos with the built-in camera of a smart-phone, and used both the raw green trace and the ICA separated sources [12]. Li et al proposed a HR measurement framework, which can reduce the noises caused by illumination variations and subjects' motions [14]. Tulyakov improved [14] by Self-Adaptive Matrix Completion [30]. Xu et al. proposed a pixel quotient in log space to derive signal for computing the pulse heart rate [28]. It is based on a model of light interaction with human skin, in which the pixel quotient in log space is robust to illumination variations. Different from them, we use a RealSense cameras with near-infrared (NIR) channel. This new camera does not depend on the environment illuminations. Thus it is more robust than the pixel quotient [28] and it can work well in almost dark conditions (see Fig. 8). On the other hand, Zheng et al. proposed to use thermal (infrared) images for HR measurement [29], where variations in temperature are successfully associated with heart beats. Though thermal

978-1-4673-8910-5/16/$31.00 ©2016 IEEE
Authorized licensed use limited to: Institute of Science and Technology Austria (IST). Downloaded on September 22,2026 at 11:13:12 UTC from IEEE Xplore. Restrictions apply.

measurement provides an illumination invariant solution, thermal imaging devices are in general very expensive.



(a) (b) (c)

Fig. 1. Three channels of the RealSense camera. (a) RGB channel; (b) Near-infrared channel; (c) Depth channel.



(a) (b)

Fig. 2. Videos captured under different illuminations

It has been reported that skin color changes caused by cardiac pulse can be captured by ordinary cameras for HR measurement [15, 23], but this is a challenging task since the change caused by the cardiac pulse is very small compared to other factors that can also cause fluctuation of the RGB value of local skin. Among these factors, illumination variation is a key one. The performance of all these methods [2, 12, 14, 15, 16] drops significantly when there are serious illumination changes. To this end, we use the very low-cost (in the order of $20) RealSense camera [31] to solve this issue.

Intel's newly-announced very low-cost ubiquitous RealSense 3D (RGBD) camera has become available very recently, being embedded into the screen lids of a dozen of laptop models from major brands such as Lenovo, Dell, and HP. RealSense is the first integrated 3D sensor to reach the consumer market, and has better short-range resolution than other low-cost platforms such as those developed by PrimeSense. RealSense produces RGB images with extra depth information (see Fig. 1), which is inferred from its latent near-infrared (NIR) channel.

Here are the contributions of our method:

- Motivated by pulse detection in RGB images, we propose to utilize the hidden near-infrared channel in the RealSense camera to enable illumination invariant pulse estimation (see Fig. 2). Near-infrared, as commonly adopted in night vision systems, provides an illumination invariant low-cost alternative. However, to the best of our knowledge, there is no clear evidence reported that near-infrared images provide a reliable source for pulse detection, as they capture very different wavelengths from both thermal and RGB images. We experimentally demonstrate, for the first time, that heart rate can be reliably estimated from faces in near-infrared images. Such observations not only enable pulse estimation to be invariant to light conditions, but also significantly extend its usage to low-light applications, such as night driving. Our results also suggest a novel way to utilize the near-infrared channel in RealSense and similar technologies, originally hidden from consumers.
- We develop a global self-similarity (GSS) filter to filter the infrared channel.
- We use depth information provided by RealSense cameras to improve the location accuracy of the region of interests (ROI) in faces (i.e., cheek regions).

II. METHOD

In this section we introduce our framework on how to use the RealSense camera for heart rate measurement.

A. Framework

The proposed framework is shown in Fig. 3. We use the RealSense camera to simultaneously record the RGB video, the NIR video, and the depth video of a person's face (see Fig. 1). We perform face detection and landmark tracking on the NIR video [27], and blob analysis on the depth video to isolate and segment the cheek region of the face in each frame (see Section II.B). We then compute the average NIR intensity of the selected region and carry out the detrending step to remove the absolute intensity variations [21] (see Fig. 5). Subsequently we apply GSS filter (see Section II.C). After this GSS filtering, we normalize the signal and use a temporal moving-average filter to remove random noise. After that, we use a band pass filter to cut outside of [0.7, 4] range in the frequency domain, which is the normal interval of heart rate of a human, i.e., [42, 240] beat-per-minute (bpm). Finally, we perform an FFT to transform the signal from the spatial domain to the frequency domain and estimate its power spectral density (PSD) distribution using Welch's method [25]. We use the frequency with the maximal power response as the HR frequency f_{HR} (Fig. 5(g)), obtaining the average HR measured from NIR input video as

$$\Psi_{HR} = 60 \times f_{HR}. \quad (1)$$

We will use the color-based method proposed in [14] to process the RGB videos for comparison. Under well-illuminated conditions, our method is able to achieve comparable results. As for situations where the illumination is low or changing, the performance of the methods using RGB images drops significantly while our method still attains the same level of performance.

B. Region selection and tracking

It has been shown in [2, 14], that the cheek and forehead areas of the face are an ideal region for heart rate estimation. We combine facial landmark tracking and depth information from the RealSense camera to automatically segment out the cheek regions. The steps are shown in Fig. 6.

Authorized licensed use limited to: Institute of Science and Technology Austria (IST). Downloaded on September 22,2026 at 11:13:12 UTC from IEEE Xplore. Restrictions apply.

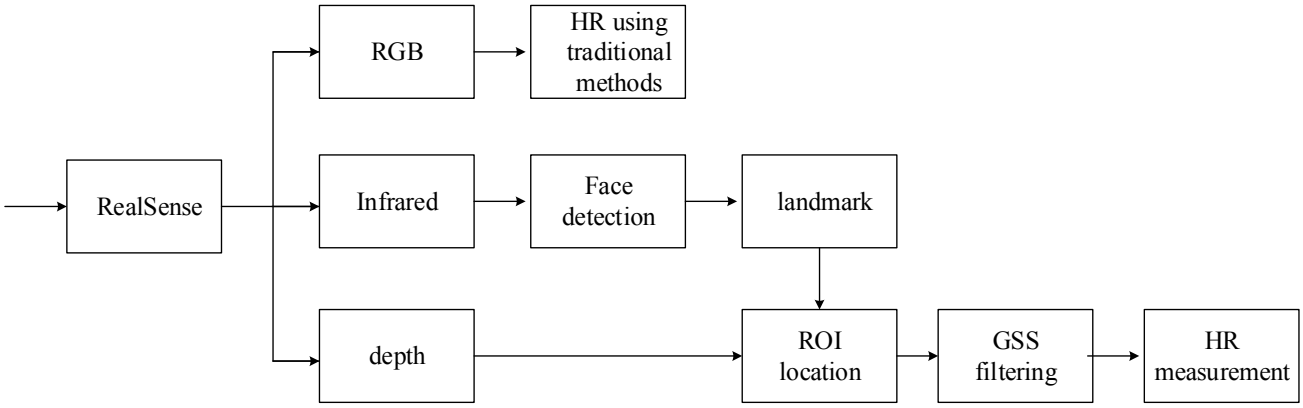


Fig. 3. Proposed framework

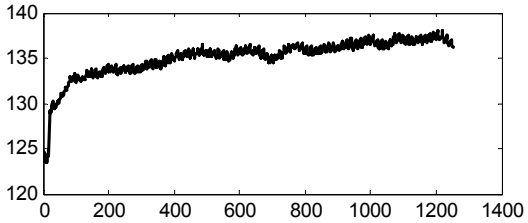
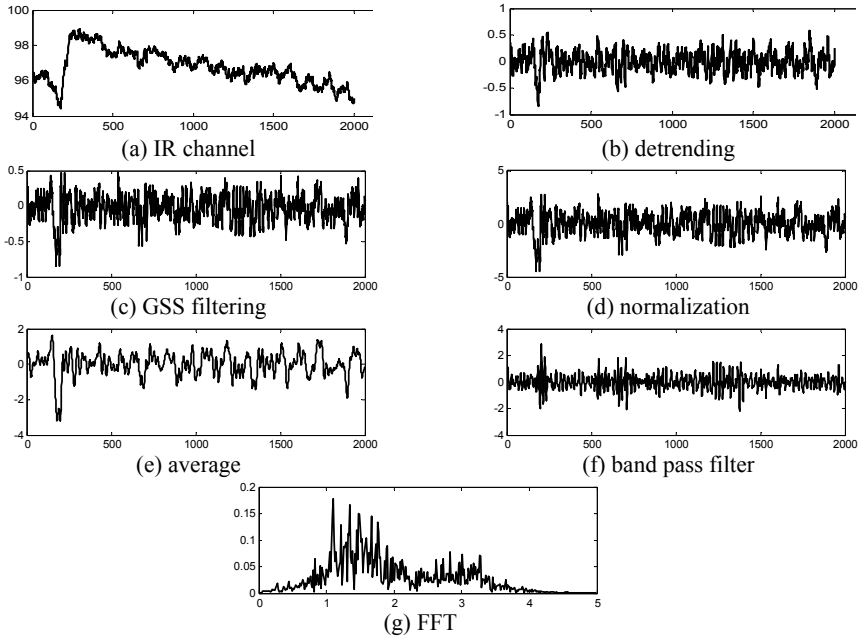
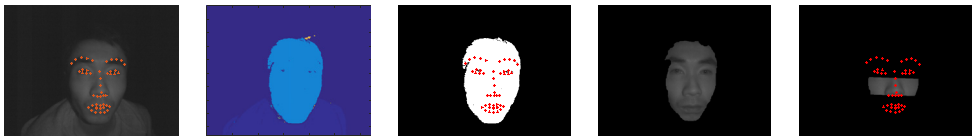


Fig. 4. Average intensity of green channel for ROI



(a) IR channel (b) detrending (c) GSS filtering (d) normalization (e) average (f) band pass filter (g) FFT

Fig. 5. Heart rate measure for ROI in infrared channel



(a) (b) (c) (d) (e)

Fig. 6. Cheek region segmentation using NIR and depth images. (a) Facial landmarks tracked on NIR image; (b) Connected components in depth image; (c) The connected component containing the most landmarks is selected as the face. (d) The face on the NIR image; (e) The cheek region is selected as area between the eyes and mouth landmarks.

Authorized licensed use limited to: Institute of Science and Technology Austria (IST). Downloaded on September 22,2026 at 11:13:12 UTC from IEEE Xplore. Restrictions apply.

We first apply the Viola-Jones face detector [24] on the infrared video to find a bounding box containing the face. A supervised method [27] is then used to track landmarks on the face (Fig 6(a)). The connected components of the depth image are computed (Fig. 6(b)). The landmarks are then mapped from the NIR image to the depth image and the connected component, excluding background, containing the majority of facial landmarks to be the face (Fig. 6(c)). Morphological operations such as erosion and hole filling are then performed on this region before using it as a mask to segment the face region in the NIR image (Fig. 6(d)). Using the landmarks, the cheek is then extracted selecting the area beneath the eyes and above the mouth (Fig. 6(e)).

C. GSS filter for infrared channel

In the infrared channel ζ_{IR} (see Fig. 5(b)), there are some noise in this signal. To filter out the noise, we propose a GSS filter to smooth the infrared channel, which improves the accuracy of the heart ratio measurement.

Self-similarity is an attractive image property which has recently found its way into object recognition in the form of local self-similarity descriptors [7]. Similarly, we find the same local self-similarity in the infrared channel (see Fig. 5(b), (c)). In addition, GSS guides the global topological structure and work better than local ones.

Specifically, for each frame, we compute the ROI as shown in Fig. 6. For the ROI in Fig. 6(e), we compute the average pixel intensity of this ROI, and denote it as ζ_{IR}. For each video, we have $\zeta_{IR} = \{\zeta_{IR,i}, i = 1,...N\}$; where N is the number of frames in the video; $\zeta_{IR,i}$ is the average intensity of the ROI of the i-th frame.

We divide ζ_{IR} of a video into T segments: $\zeta_{IR} = \{\zeta_\tau, \tau = 1,2,...T\}$. Each segment corresponds to a length of 1.5 seconds. We then perform clustering for all T segments ζ_τ. After clustering, we have K clusters $C = \{c_1, c_2, ...c_K\}$ and Δ_{sumd}, where Δ_{sumd} is the within-cluster sums of point-to-centroid distances. We discard the m clusters with the largest Δ_{sumd}, i.e., $C' = \{c_{K-m}, ...c_K\}$, and keep those clusters which have smaller Δ_{sumd}, i.e., $C^- = \{c_1, ...c_{K-m-1}\}$. For those segments belong to C', we use the neighbouring previous segment to replace them. For example, if $\zeta_\tau \in C'$, we use $\zeta_{\tau-1}$ to replace it. Here $\zeta_{\tau-1}$ should be in C^-. Otherwise, it has been replaced by its neighbouring pervious segment.

In our case we let K =10, m=2. These two parameters are used to control the smoothing of the filtering. If K becomes larger, we have less segments in each cluster and the smoothing is marginal, and vice versa. If m becomes larger, we discard more clusters and the smoothing becomes stronger, and vice versa.

The collected dataset in our case is usually around 90 seconds. We have 60 segments for one video. For each segment, we use 1.5 seconds since the lower limit of normal heart rate is 42 (see Section II.A). Thus, each 1.5 seconds segment includes one heart beat (i.e., at least one period of heart beating signal).

III. EXPERIMENTAL RESULTS

We test our method for datasets collected using the RealSense camera and also compare with existing methods.

A. Data collection

We have collected two datasets (to be made publically available upon publication). The first dataset (verification dataset) is a simple dataset for verification purposes since we re-implemented previously proposed methods. The second dataset (challenging dataset) is a more challenging dataset to test the robustness of our proposed method.

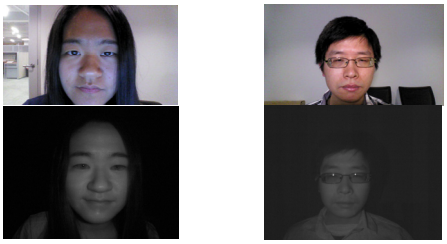


Fig. 7. Faces in frontal pose and under constrained illuminations

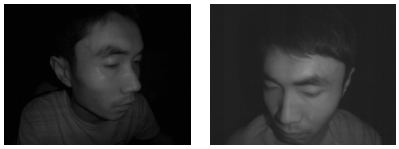


Fig. 8. Faces in multi-poses and under unconstrained illuminations

The verification dataset is collected using the RealSense camera under constrained illuminations and frontal pose. Some examples are shown in Fig. 7. Each sequence is saved in png format with different frame rates, 12 frames per second (fps), 15 fps and 30 fps. Three subjects (one females and two males) aged from 20 to 40 years old were enrolled. During the recording, subjects were asked to sit still in a chair and try to avoid any movement. Here, the subject are still during data collection, which is to follow the setup as that in [2, 16] to verify the re-implemented previously proposed methods [2, 16]. The RealSense camera was fixed on a laptop about 30 cm from the subject's face. Each subject was recorded for about 90 seconds. There are three channels for each subject, i.e., RGB, infrared and depth channels. The RGB channel is in 24-bit color format with resolution of 1920×1080. The infrared channel is in 8-bit gray format with resolution of 640×480. The depth channel is also with resolution of 640×480.

The challenging dataset is also collected using the RealSense camera but under unconstrained illuminations, multi-poses and multi-races. Some examples are shown in Fig. 8. Each sequence is saved in png format with a frame rate of 30 fps. Ten subjects (1 female and 9 males) aged from 20 to 50 years old were enrolled. Subjects were asked to sit on a chair and show different poses. The RealSense was fixed on a laptop

Authorized licensed use limited to: Institute of Science and Technology Austria (IST). Downloaded on September 22,2026 at 11:13:12 UTC from IEEE Xplore. Restrictions apply.

and about 30 cm from the subject's face. Each subject was recorded for about 90 seconds. Two channels were recorded for each subject, i.e., infrared and depth channel. Both the infrared and depth channels are in 8-bit format with resolution of 640×480.

We re-implement three previous methods: two color based ones Li 2014 [14], Poh 2011[16], and one motion-based one Balakrishnan 2013 [2]. In [2] they also used customized peak detection functions to find the location of each heart beat for further HR variation analysis. We did not replicate the peak detection process here since we only aim to compare the accuracy of the methods on estimating the average HR. The FFT is applied at the last stage for each method to find the average pulse frequency.

For these two dataset, we use two tools to measure the ground truth. One is a finger pulse Oximeter CMS50F, which gives an average heart rate for a sequence. The other one is an ECG BioRadio, REF BR-500.The sampling rate is 250HZ. It detects the heart rate beat by beat as shown in Fig. 9

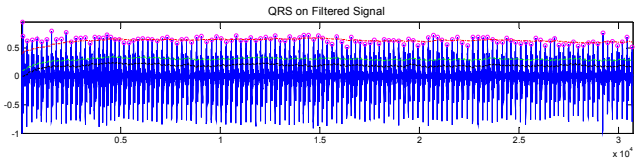


Fig. 9. Hear rate measurement beat by beat using BioRadio

B. Results for verification dataset

The results of all methods for the verification dataset are shown in Tables 1-4. In Tables 1 and 2, we show the HR result for each sequence for both RGB and infrared channels. In tables 3 and 4, we give the statistical results. Here, the measure error is computed as

$$\Psi_{error} = \Psi_{HR} - \Psi_{GT}, \tag{2}$$

where Ψ_{HR} denotes HR measured from video (see Eq. (1)), and Ψ_{GT} is the ground truth HR.

To comprehensively compare the methods in multiple aspects, we include five kinds of statistics used in previous research works. The first one is the mean of Ψ_{error} denoted as M_e; the second one is the standard deviation of Ψ_{error} denoted as SD_e; the third one is the root mean squared error denoted as *RMSE*; the fourth one is the mean of error-rate percentage

$$M_{eRate} = \frac{1}{N}\sum_{l=1}^{N}\frac{|\Psi_{error}(l)|}{\Psi_{GT}} \tag{3}$$

where N is the number of videos in the database.

The fifth is accuracy, computed as

$$Accuracy = \frac{N_{correct}}{N}, \tag{4}$$

where $N_{correct}$ is the number of videos which have the correct HR measurement. The HR measurement for one video is correct if $|\Psi_{error}| \le \Gamma$. Here the threshold is Γ=5, suggested by a consulted medical specialist in the team.

Table 1: Illustration results for RGB channel of verification dataset (GT: ground truth)

Method	(a)  (GT: 62.99)	(b)  (GT: 69.1)	(c)  (GT: 69.1)
Poh 2011	60.35	68.99	68.55
Balakrishnan 2013	62.89	65.51	68.66
Li 2014	60.13	67.54	68.21
Proposed	60.35	67.67	68.55

Table 2: Illustration results for infrared channel of verification dataset

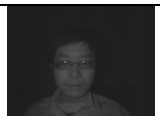
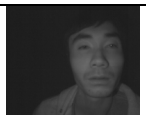
Method	(a)  (GT: 62.99)	(b)  (GT: 69.1)	(c)  (GT: 69.1)
Poh 2011	60.64	66.53	57.12
Balakrishnan 2013	61.34	136.41	69.58
Li 2014	60.26	67.18	68.23
Proposed	60.44	67.23	68.55

Table 3: Statistical results for RGB channel of verification dataset

Method	$M_e(SD_e)$ (bpm)	RMSE (bpm)	M_{eRate} (%)	Accuracy (%)
Poh 2011	1.10(1.91)	1.10	0.0172	100
Balakrishnan 2013	1.37(2.72)	1.37	0.0200	100
Li 2014	1.77(1.41)	1.77	0.0270	100
Proposed	1.54(1.48)	1.54	0.0235	100

Table 4: Statistical results for infrared channel of verification dataset

Method	$M_e(SD_e)$ (bpm)	RMSE (bpm)	M_{eRate} (%)	Accuracy (%)
Poh 2011	5.63(7.77)	5.63	0.0826	66.7
Balakrishnan 2013	-22.04(55.45)	23.14	0.3357	66.7
Li 2014	1.84(1.31)	1.84	0.0279	100
Proposed	1.65(1.43)	1.65	0.0252	100

From Table 1, we can see that all methods work very well for high resolution images (1920×1080) under constrained conditions, i.e., frontal pose and indoor illuminations.

From Table 2, we can see that both our method and Li 2014 work very well. Poh 2011 perform well for the first two sequences but fails for the third one. One explanation might be due to the shadow around the nose. Balakrishnan 2013 correctly measures the HR for the first and third sequences but not the second one. The reason is that the face in the middle sequence is small and a lot of tracking points move away from the region of interest during the last few frames.

Authorized licensed use limited to: Institute of Science and Technology Austria (IST). Downloaded on September 22,2026 at 11:13:12 UTC from IEEE Xplore. Restrictions apply.

Tables 3 and 4 further show the robustness of our proposed method to illumination changes, thanks to the use of the NIR and depth channels. As we will further show next, our method is the only one robust both to illumination and pose.

C. Results for challenging dataset

Table 5: Results for infrared channel of challenging dataset

Method	$M_e(SD_e)$ (bpm)	RMSE (bpm)	M_{eRate} (%)	Accuracy (%)
Poh 2011	-8.40(27.98)	15.04	0.2259	70
Balakrishnan 2013	-5.91(17.95)	10.33	0.1507	90
Li 2014	-1.45(7.99)	4.56	0.0632	90
Proposed	2.26(6.54)	3.66	0.0534	100

The results for the challenging dataset are shown in Table 5. Clearly the proposed method works the best. It correctly measures the heart rate of all the videos. Both Li 2014 and Balakrishnan 2013 perform well for this dataset and measure 90% of the videos correctly. Poh 2011 also works well for this challenging dataset, although pose variations degrade the performance of this method. Although this dataset show serious illumination variations, the video modalities captured using the RealSense camera addresses this issue with our proposed method.

CONCLUSION

In this paper we demonstrated the use of the low-cost Intel RealSense RGBD camera to measure heart rate. Simultaneously exploiting the depth and NIR information, we proposed a method that is invariant to illumination and face pose, therefore opening the application of touch-less hear rate monitoring to new scenarios such as night driving and monitoring in the wild.

ACKNOWLEDGEMENT

This work was sponsored by the Academy of Finland, Infotech Oulu and partially supported by ONR, ARO, NSF and NGA.

REFERENCES

[1] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic. Robust discriminative response map fitting with constrained local models. In *CVPR*, 2013.

[2] G. Balakrishnan, F. Durand, and J. Guttag. Detecting pulse from head motions in video. In *CVPR*, 2013.

[3] R. Basri and D.W. Jacobs. Lambertian reflectance and linear subspaces. *TPAMI*, 2003.

[4] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[5] G. Cennini, J. Arguel, K. Akşit, and A. van Leest. Heart rate monitoring via remote photoplethysmography with motion artifacts reduction. *Optics express*, 2010.

[6] K. Chan and Y. Zhang. Adaptive reduction of motion artifact from photoplethysmographic recordings using a variable step-size lms filter. *In Proceedings of IEEE on Sensors*, 2002.

[7] T. Deselaers and V. Ferrari, Global and Efficient Self-Similarity for Object Classification and Detection, *CVPR* 2010

[8] F. X. Gamelin, S. Berthoin, and L. Bosquet. Validity of the polar s810 heart rate monitor to measure rr intervals at rest. *Medicine and Science in Sports and Exercise*, 2006.

[9] M. H. Hayes. 9.4: Recursive least squares. *Statistical Digital Signal Processing and Modeling*, 1996.

[10] K. HumpPAeys, T. Ward, and C. Markham. Noncontact simultaneous dual wavelength photoplethysmography: a further step toward noncontact pulse oximetry. *Review of scientific instruments*, 2007.

[11] V. Jeanne, M. Asselman, B. den Brinker, and M. Bulut, Camera-based heart rate monitoring in highly dynamic light conditions, *International Conference on Connected Vehicles and Expo*, 2013

[12] S. Kwon, H. Kim, and K. S. Park. Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. In *EMBS*, 2012.

[13] C. Li, C. Xu, and et al. Distance regularized level set evolution and its application to image segmentation. *IEEE Trans. on Image Processing*, 2010.

[14] X. Li, J. Chen, G. Zhao and M. Pietikäinen, Remote heart rate measurement from face videos under realistic situations. *CVPR* 2014

[15] M.-Z. Poh, D. J. McDuff, and R. W. Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 2010.

[16] M.-Z. Poh, D. J. McDuff, and R. W. Picard. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Trans. on Biomedical Engineering*, 2011.

[17] S. Prahl. Optical absorption of hemoglobin. http:// omlc.ogi.edu/spectra/hemoglobin/, 1999.

[18] J. Shi and C. Tomasi. Good features to track. In *CVPR*, 1994.

[19] H. Simon. *Adaptive filter theory*. Prentice Hall, 2002.

[20] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic. A multimodal database for affect recognition and implicit tagging. *IEEE Trans. on Affective Computing*, 2012.

[21] M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen. An advanced detrending method with application to PAv analysis. *IEEE Trans. on Biomed. Eng.*, 2002.

[22] C. Tomasi and T. Kanade. *Detection and tracking of point features*. CMU, 1991.

[23] W. Verkruysse, L. O. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Optics express*, 2008.

[24] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *CVPR*, 2001.

[25] 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*, 1967.

[26] H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, W. Freeman, Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph*. 2012

[27] X. Xiong and F. De la Torre, Supervised Descent Method and its Applications to Face Alignment, *CVPR* 2013

[28] S. Xu, L. Sun, and G. K. Rohde, Robust efficient estimation of heart rate pulse from video, *Biomedical Optics Express*, 2014

[29] B. S. Zheng, M Murugappan, S. Yaacob, Human Emotional Stress Assessment through Heart Rate Detection in a Customized Protocol Experiment, *IEEE Symposium on Industrial Electronics and Applications*, 2012.

[30] S. Tulyakov, X. Alameda-Pineda, E. Ricci, L. Yin, J. F. Cohn, Nicu Sebe, Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions, CVPR 2016

[31] http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-3d-camera.html

Authorized licensed use limited to: Institute of Science and Technology Austria (IST). Downloaded on September 22,2026 at 11:13:12 UTC from IEEE Xplore. Restrictions apply.