

Let home nursing assistant robots see your heart rate

Han Wu, Tao Wang, Tuo Dai, Xiaoyu Wang, Yuanzhen Lin and
Yizhou Wang

*School of Electronic Engineering and Computer Science,
Peking University, Beijing, China*

Abstract

Purpose – This paper aims to design a vision-based non-contact real-time accurate heart rate (HR) measurement framework for home nursing assistant.

Design/methodology/approach – The study applied Second-Order Blind Signal Identification (SOBI) algorithm to extract remote HR signal and analyzed it with Fast Fourier Transform (FFT). Multiple regions of interest are chosen and analyzed to obtain a more accurate result.

Findings – An accurate non-contact heart rate (HR) measurement framework is proposed and proved to be efficient.

Originality/value – The contributions of this HR measurement framework are as follows: accurate measurement of HR, real-time performance, robust under various scenes such as conversation, lightweight computation which is suitable and necessary for home nursing assistance. This framework is designed to be flexibly used in various real-life scenes such as domestic health assistance and affectively intelligent agents and is proved to be robust under such scenes.

Keywords Crowd AI, Crowdsourcing human-robot interaction

Paper type Research paper

1. Introduction

Heart rate (HR) is an important physiological index reflecting both health condition and emotional state. As the American Heart Association states, the normal resting human HR is a range between 60 and 100 beats per minute (bpm) (AHA, 2017). Usually, activities like physical exercise, sleep, illness and emotional swings like anxiety, stress and excitement can result in HR changes. Too fast and too slow HR can happen at certain occasions such as exercise and sleep. But irregular patterns, abnormal HR and mutations can indicate diseases. Consistently monitoring HR can obviously play an important role in keeping track of the health condition of the elderly. When it comes to home nursing assistant robots, this monitoring process requires not only accuracy but also user-friendliness, which means causing as less disturbance as possible. Under such circumstances, vision-based non-contact methods are preferred to other measurement methods.



Vision-based HR measurement has been studied for many years. Most of the researches are based on remote photoplethysmogram (rPPG). rPPG-based HR measurement uses the change of light reflected by blood flow to detect heartbeats and various algorithms have been proposed to cope with this problem. There are two main components of HR measurement algorithms, rPPG signal extraction and HR calculation. Traditional HR measurements are usually based on signal processing and analyzing methods such as blind signal separation (BSS) and Fast Fourier Transformation (FFT). As data-driven methods arousing more and more attention, machine learning and deep learning have been applied to this task and obtained good performance on datasets.

As mentioned above, HR measurement for home nursing assistance requires not only the accurate HR value, but also timeliness and user-friendliness. Such HR measurement frameworks should be capable of measuring elder people's HR when interacting with them in daily scenes to keep aware of their physical health and emotional states. Thus, a practical HR measurement for home nursing assistance needs to detect human HR accurately without physical contact and be robust enough to allow slight body shaking which is normal during conversation. Most of the previous works focused on modifying and examining their methods on recorded videos while sparsely investigating its practicability in real-world scenario and hardly pragmatic for home nursing assistance. It should also be low cost as the assistant robot need to move around in the room which rules out most of the data-driven methods because they usually require an external graphics processing unit (GPU).

In this paper, we propose a real-time HR measurement framework that can detect human HR from normal webcams. Our framework applies second-order blind signal identification (SOBI) algorithm to tackle the BSS problem and selects the rPPG signal based on spectral kurtosis, which is then verified by the spectral power distribution. As we focus on HR measurement methods suitable for home nursing assistance, it requires rapid and accurate HR measurement instead of other features related to amplitudes. Under such motivation, a series of glitch removal methods and a peak detection method are employed to extract HR from the raw rPPG signal.

The contributions of this HR measurement framework are as follows:

- accurate measurement of HR;
- real-time performance;
- robust under various scenes such as conversation; and
- light-weight computation which is suitable and necessary for home nursing assistance.

This framework is designed to be flexibly used in various real-life scenes like domestic health assistance and affectively intelligent agents and is proved to be robust under such scenes.

The remainder of this paper is organized as follows. Section 2 gives an overview of the related works in the field of rPPG-based HR measurements. Section 3 explains main challenges in designing such a framework and our solutions to them. Section 4 details the design of the proposed real-time vision-based HR measurement framework. In Section 5, we evaluate the proposed framework and compare it with the ground truth obtained from a home pulse oximeter, followed by conclusion and upcoming works in Section 6.

2. Related works

Generally, rPPG-based HR measurement methods use a sequence of regions of interest (ROIs) in human face obtaining from various sensors to extract its hidden rPPG-signal

and analyze it. Unavoidably, the observed signal usually contains not only the clean rPPG signal we wish to analyze but also the noise from environment such as light variation, thermal noise, power frequency interference and other uncatalogued noises. One important procedure of rPPG-based HR measurement methods is the noise removal and clean rPPG signal separation. Most of the existing methods adopt traditional BSS algorithms. BSS is the separation of a set of source signals from a set of mixed signals, without the aid of information (or with very little information) about the source signals or the mixing process (Wiki, 2018). Some of the famous BSS algorithms are independent component analysis (ICA), principal component analysis (PCA), singular value decomposition (SVD) and so on.

Poh *et al.* (2011) put forward a method which employs a JADE implementation of ICA and takes the three channels of ROI sequence as the input to JADE. They recorded video streams to evaluate their method and extended this methodology to obtain more physiological indexes like respiratory rate (RR). Their work proved that at the sampling rate of a webcam, the data collection is sufficient for rough rPPG signal obtaining. Kwon *et al.* (2012) developed an iOS application which is also based on video recording and ICA analysis, but instead of all three channels of ROI sequence, they only used the data from green channel. These two methods are based on video recording and yet unknown for their real-time performance. Other than ICA-based methods, Jiang *et al.* (2014) creatively used green channel data as input to a Kalman Filter to enhance the signal and developed an Android application to estimate the proposed algorithm. Their method is proved to have a closer value to the ground truth compared to ICA. This real-time HR measurement application needs users to keep their face in the Region of Interest which is unfriendly and not practical to use in human-computer interaction for the elderly.

Machine learning has also showed its potential in HR measurement. Kessler *et al.* (2016) used k-Nearest Neighbor (kNN) and multi-layer perceptron (MLP) with an alternative representation of the input vector to learn the regression of HR and the data sequence from green channel. They improved the root mean square error (RMSE) from 23.97 to 8.62, which successfully ameliorated the accuracy of HR measurement to some extent. Some researchers also tried to combine signal processing methods to machine learning methods. By combining joint blind signal analysis (JBSS) algorithm and machine learning, Qi *et al.* (2017) proved their proposed method outperformed traditional ICA-based methods on a data set and achieved good accuracy. As a matter of fact, machine learning has been used in BSS for a long time. Wei *et al.* (2007) successfully recovered the source signals from a set of nonlinear underdetermined mixed signals by combining Bayesian statistics with MLP neural network. However, this kind of BSS methods lacks the balance between accuracy and computation load and is not practical for real-time applications.

Compared to the conference paper of our precedent work published in ICAA2018 (Wu *et al.*, 2018), this paper put more effort in the multiple ROIs fusion part which will be explained in Section 3. We also supplemented our experiment to validate this method.

3. Challenges and solutions

One of the difficulties of designing a HR measurement framework suitable for home nursing assistant robots lies in balancing accuracy and strong timeliness. In real-world scenario, the sampling rate even of a normal webcam can be up to hundreds of frames per second (fps) and according to the sampling theory, the sampling rate should be at least 15 to 30 fps for a HR measurement. Thus it puts forward a challenge for employing an appropriate algorithm that can process the sample data with low-cost computation capability in pace with the sampling speed.

Otherwise the sample data will be piled up and as time goes on and the framework will gradually lose its real-time property. Another challenge of HR measurement under such circumstances has been troublesome and worthy studying. When the user is not required to stay still, the collected ROI data is usually contaminated severely and hardly can be used directly as the input of a BSS problem. But in the application of home nursing assistant robot, it is necessary not to require users to stay still, which brings much trouble to the design of the framework.

In this paper, we tackle these two challenges by several ways. First, low-cost but highly efficient signal processing methods such as SOBI and spectral kurtosis are used which are accurate and yield no burden on the computation. Second, the design of the framework is devoted to balancing the sample collection rate and the calculation speed. For example, in the proposed framework, the ROI detection method is not applied to every frame as people are relatively still when interacting with the nursing agents instead of shaking fiercely, which helps to reduce the computation load while its accuracy is still guaranteed. Last but not least, we discovered that the rPPG signal data retrieved from forehead area and cheeks are almost of no delay, which is obvious in Figure 1. So the supplement of data from several ROIs can prevent data loss when the user is at specific pose such as side facing the camera. We calculate the correlation coefficient between signals gained from different ROIs to decide whether signals can be adding up to strength the intensity or one of the potential signals should be chosen. Under the later circumstance, the target signal is chosen based on the signal frequency distribution.

4. Proposed framework

The main procedure of our proposed framework consists of five steps as Figure 2 shows. First, a boosted cascade of classifiers based on Haar-like features (Viola and Jones, 2001) is applied to locate human face in the video stream and a face landmark detection method from Dlib (King, 2009) is used to locate key points in the face and calculate the ROI coordinates. After that, the data from all the three RGB channels of ROI are collected and spatially normalized over all pixels. We also apply a Butterworth filter (Wiki, 2018) to remove signal components whose frequencies are lower than 0.5 Hz (30 beats per minute) or higher than 4

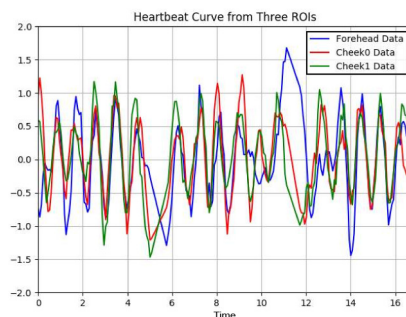


Figure 1.
Heartbeat curve of
data from different
ROIs



Figure 2.
Structure of the
proposed framework

Hz (240 beats per minute) which are nearly impossible for resting adults. Once enough ROI data have been collected for analysis, they are fed into the SOBI method to extract the hidden rPPG signal. The output of SOBI is composed of three independent signals, one of which is the expected rPPG signal. The output sequence of the three signals is at random. Thus, a signal selection method is needed to pick out the rPPG signal. The signal selection is based on spectral kurtosis and verified by spectral power distribution to ensure the correct signal is to be analyzed. The last step is to calculate the HR. To improved accuracy, the rPPG signal is first smoothed by a shifting window filter and then analyzed by Fast Fourier Transform (FFT). The HR value is calculated based on the properties of rPPG signals extracted from different ROIs. At the same time, a peak detection method is applied to draw the heart beat curve simultaneously.

4.1 Face detection and face landmark detection

We applied a boosted cascade of classifiers based on Haar-like features (Viola and Jones, 2001) to detect human face. This widely used classifier shows advantages in processing rapidly and getting relatively accurate results enough for ROI location, which helps accelerate the calculation speed. The detected face is then resized and used as the input for the landmark detector from Dlib. (King, 2009)

The motion artifacts are obvious interference to HR calculation. Even if the signal is taken from sensors placed on human body, motion artifacts can still result in much disturbance and pollute the clean signal, which has been proved and studied by many researchers (Elgendi, 2012; Lee and Zhang, 2003). Instead of studying how to remove motion artifacts, we compared and chose three facial areas which are the least affected by head motion and have strong signal strength.

The forehead and cheek areas are rich in capillaries which lead to stronger rPPG signal than other regions. In the meantime, these regions are almost unaffected by facial expressions especially the forehead area. The former rPPG-based HR measurements usually choose a large percentage of human face as ROI including the eyes and lips which can result in much motion artifacts and in the end, a less accurate estimation of HR. Also, the idea of choosing dual ROIs can help to verify the accuracy and supplement when necessary.

Therefore, in this framework, three small rectangles in the middle of forehead and two cheeks are chosen as the Regions of Interest (ROIs). Considering the real-world scenario when people interact with nursing agents, it is not practical or friendly to ask users to stay absolutely still. This is why a steady face landmark detection algorithm is used here. The coordinate of forehead ROI is calculated by locating the eyebrow tip and cheek ROIs is determined by the corners of mouth. Besides, by specifying the size of ROI in every frame, it is guaranteed that all the samples are taken from the same region with the same size.

4.2 Regions of interest data collection and preprocessor

In the calculation of HR, the rPPG signal needs enough sampling points to be accurate. According to European cardiology task force, the optimal range for HR analysis is 250 to 500 Hz or perhaps even higher to precisely recover the details of HR information (Dwyer, 1984). In fact and in common sense, the higher the sampling rate is, the more accurate the HR measurement result will be. In an rPPG-based program, the sampling rate is usually limited by the property of webcams and when it comes to real-time scenario, the computation capability also sets a boundary for the sampling rate to some extent. Under the motivation of accurate and rapid HR measurement, the ROI data

collection is consistent through the whole process, which is easy to achieve in the scene of human–computer interaction. Thus the delay of HR measurement can be greatly reduced.

If no data are available at first, then in the very beginning of the HR measurement only 5 s are needed to accumulate enough ROI data and in the subsequent analysis, the ROI data collection and HR calculation will be conducted at the same time. Data from the three ROIs are analyzed in the same way respectively before the HR calculation part.

Every valid ROI is separated into three RGB channels. The data in each channel are spatially averaged to yield one sampling point. Combined with the sampling time, these three channels will then form the three raw signals $r_1(t)$, $r_2(t)$, $r_3(t)$. In the raw signals, there are some frequency components lower than 0.75 Hz or higher than 4 Hz which are irrelevant for the purpose of HR calculation. A Butterworth filter is applied here to filter out these frequencies. Then a z-score is placed on r_i to standardize it as follows:

$$r'_i = \frac{r_i - \mu_i}{\sigma_i} \quad (1)$$

For each $i = 1, 2, 3$, μ_i and σ_i are the mean value and standard deviation of r_i . After being preprocessed in the above ways, three normalized signals containing rPPG information are prepared.

4.3 rPPG signal decomposition

The three raw signals r'_i , $i = 1, 2, 3$ contains the expected rPPG signal, and thus are supposed to be decomposed into three independent signals based on the second-order blind signal identification method. We run several tests based on other methods such as fastICA, wavelets and RNN and results showed that SOBI outperforms others in rPPG signal decomposition.

The explanation to SOBI is as follows. Given an observed signal $\chi(t)$, it is formed by n signals, in our case where n is 3. Each of $\chi_i(t)$ can be considered as a linear instantaneous mixing of n source signals $s_i(t)$ by a mixing matrix A , which means $\chi(t) = A \cdot s(t)$. What SOBI can do is estimating a decomposition matrix W similar to A based on the observed signal $\chi(t)$. Thereby, source signals $s'(t)$ can be estimated under (2).

$$s'(t) = W\chi(t) = s(t) \quad (2)$$

The estimation of composition matrix is based on matrix diagonalization. The first step is to construct a set of diagonal matrices under (3) by choosing a set of time delay τ and calculating the symmetric correlation matrix of $\chi(t)$ and $\chi(t + \tau)$.

$$R_\tau = \text{sym}(\langle \chi(t)\chi(t + \tau)^T \rangle) \quad (3)$$

Where $\text{sym}(M) = \frac{M + M^T}{2}$, $\langle \rangle$ calculates the mean value over time domain.

The next step is to minimize (4) by rotating matrix V and iteration.

$$\sum_{\tau} \sum_{i \neq j} (V^T R_\tau V)_{ij}^2 \quad (4)$$

Then the decomposition matrix W can be estimated by $W = V^T B$, where $B = \text{diag}\left(\lambda_i^{-\frac{1}{2}}\right) U^T$. λ_i is the eigenvalue of correlation matrix $\langle \chi(t) \chi(t + \tau)^T \rangle$ and U is its corresponding eigenvector. $S'(t)$ is then computed based on the estimated matrix W .

The output of SOBI is three independent signals, one of which is the hidden rPPG signal. The output order is at random so an rPPG signal selection and verification method is needed. In the proposed framework, the rPPG signal is selected based on spectral kurtosis (SK) and then verified by spectral power distribution. Spectral kurtosis is defined as the kurtosis of a signal's frequency components. It was proposed to detect randomly occurring signals (Dwyer, 1984). It is now commonly used to indicate the presence of series of transients in the frequency domain. By accumulating periodic transients, period signals like rPPG signal can be distinguished by its spectral kurtosis which is obviously larger than that of non-period signals. In the proposed framework, the spectral kurtosis value of all the three independent signals are calculated by the following equation:

$$SK(z) = \frac{E(z^4)}{E(z^2)^2} - 3 \quad (5)$$

where z is one of the independent signals and z^k stands for its k th order cumulant and $E(z^k)$ can be seen as the average of z^k over time domain. The rPPG signal is thus selected based on its SK value. The SK value calculation can be done rapidly which ensures the timeliness of this real-time framework.

4.4 Heart rate calculation

The chosen rPPG signal is then smoothed by a shifting window filter with the length of 5. The intention here is to eliminate the glitches and prepare the rPPG signal for peak detection. From each ROI, a rPPG signal is prepared. The correlation coefficient value between every two rPPG signals from different ROIs is calculated to decide whether these two signals can be added up to strengthen the signal intensity. If there are no signals with high similarity, then we evaluate each signal using its frequency distribution by equation (6).

$$\text{rank}(s) = \frac{\sum_{i=k-2}^{i=k+2} a_i^2}{\sum_{i=0}^{i=n} a_i^2} \quad (6)$$

where s is a rPPG signal and $\text{rank}(s)$ is its evaluation value. The variation k stands for the frequency with the highest amplitude and a_i is the amplitude of frequency i . n is the total number of frequencies. Fast Fourier Transform (FFT) is performed on the target rPPG signal (strengthened or selected). The frequency with the highest amplitude is the HR. Besides, a custom peak detection method is developed to calculate the number of heart beats to verify the HR and draw the heart beat curve simultaneously.

5. Evaluation

5.1 Experiment setup

All the evaluation tests are performed on a PC with an intel i7-7700K processor without using any GPU. The webcam used in these test is a normal Logitech c270.

5.2 Real-world evaluation

The real-world evaluation is designed to testify the robustness and timeliness of the framework because disturbance like body swing, head motions are quite common for home nursing assistant robots. A qualified framework should function well under such circumstances. We invited dozens of volunteers to participate in our evaluations. The evaluation is composed of three parts, each of which lasts 16 seconds. In Test 1, all the volunteers are required to sit quietly and keep still. This test is the basic one to verify the accuracy and validate the correctness of the framework. While in Test 2, volunteers can sit casually with their heads nodding or shaking normally as if they were in the interaction with the nursing assistant robots. This test is designed to simulate the interaction scenario where the nursing assistant robots have to detect human HR while they are at ease. In the last test, volunteers can speak, smile which results in facial muscle movements.

We can see some of representative results in the above figures. Figure 3 shows the extracted heartbeat curve of a volunteer when he/she is sitting quietly without noticeable movements. In Figure 4, two volunteers were asked to shake or nod their heads and their heartbeats were recorded. The left heartbeat belongs to a volunteer with very slightly head

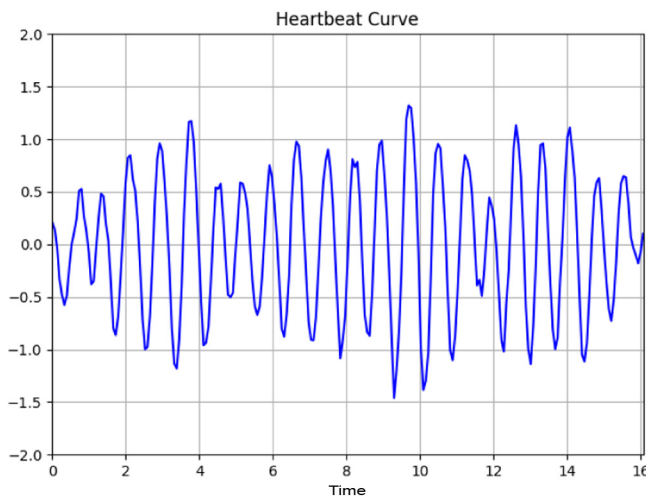


Figure 3.
Heartbeat curve in
Test 1

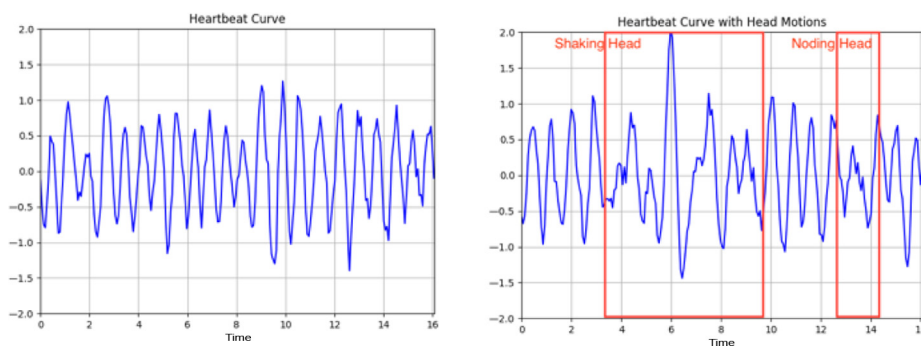


Figure 4.
Heartbeat curve in
Test 2

movements, and it can be seen that his rPPG signal is nearly under no interference, and his heartbeat curve is recovered perfectly while the right one's is disturbed by head motions but still preserves enough information to recover a complete heartbeat curve. Figure 5 shows two interlocutors' heartbeat curve recorded during their conversation. They are not as perfect as Figure 3 but still can produce accurate results of one's HR information. One of the produced real-time heartbeat curve can be seen in Figure 6.

Some of the evaluation results are shown in Table I. All these data are selected at random from test results of different volunteers and different test types. The ground truth is taken from a home pulse oximeter at the same time as the test runs. Generally, the proposed framework can almost reach an accuracy with relative error less than 1 (the biggest error is +1.2 which is shown

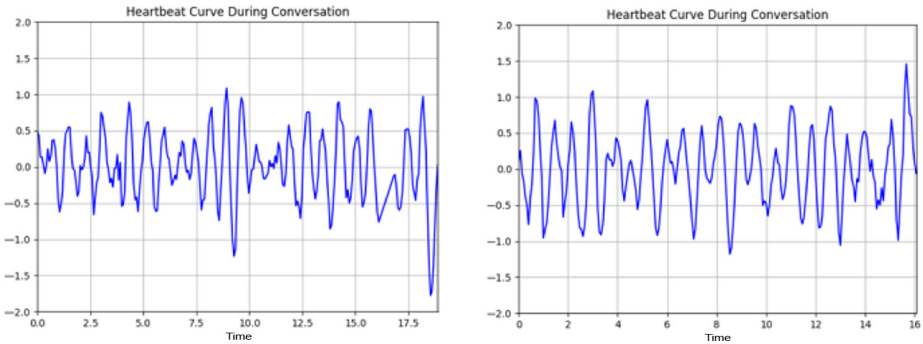


Figure 5.
Heartbeat curve in
Test 3

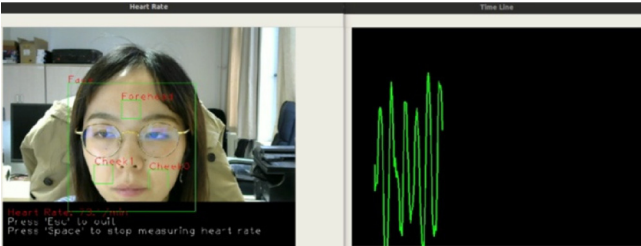


Figure 6.
A screenshot of the
real-time heartbeat
curve

Table I.
Real-world
evaluation result

Beats per minute	Experimental result	Ground truth	Deviation
Object 1	86.7	86	+0.7
Object 2	82.3	82	+0.3
Object 3	60.5	61	-0.5
Object 4	67.1	68	-0.9
Object 5	72	73	-1
Object 6	82.9	84	-1.1
Object 7	74.2	75	-0.8
Object 8	67.4	67	+0.4
Object 9	73	74	-1
Object 10	61.2	60	+1.2

in the table). The testing results shown below are selected at random and to show the effect of multiple ROIs fusion method, Object 9 and 10 are specially selected here. During these two experiments, volunteers are either talking with other people or using cellphones casually.

By evaluating this proposed framework in real-world scenario and simulate simple interaction scenes such as conversations and head motions, it is proved that this framework is accurate and robust. Also, the maximum delay of first calculation output is within 5 s which partly depends on the facial image condition (whether the user is in the frame and the pose of user *and etc.*) and fps limit of the webcam. In the following calculation, the update speed of HR measurement can be within 3 s though longer interval usually brings more accurate results.

5.3 Data set evaluation for signal processing accuracy

To validate the accuracy of signal processing, we evaluated the proposed framework on Synthetic Data set (Charlton *et al.*, 2016). This data set contains clean PPG signals from 192 objects and its HR information. We mixed random noises to the clean signal to conduct observed signal and use it as the input of the framework to verify the raw signal processing accuracy of the framework as Figure 7 shows.

The generated signals are fed into our framework to be preprocessed and analyzed. The only difference between data set evaluation and real-world evaluation is the signal obtaining way. To generate signals closer to observed ones, we carried out two tests by mixing Gaussian noise and random noise respectively. In each test, every clean PPG signal is contaminated by artificially generated noises which cannot be reproduced to ensure the verisimilitude. Because of length limit of this paper, we display the analysis result of Gaussian noise contaminated rPPG signals here.

As in Figure 8, the left picture shows one of the clean PPG signals obtained from the data set and by mixing random noise to it, we generated a contaminated signal in the right picture which is similar to observed signals in real-world scenes and unable to distinguish its cycles directly. This generated signal is processed in the proposed framework and successfully recovered as Figure 9 shows. We can see in Figure 9 that although the recovered signal is not the very same as the original one but it keeps all the key information especially the signal frequency. We carried out such experiment on all the 192 objects and

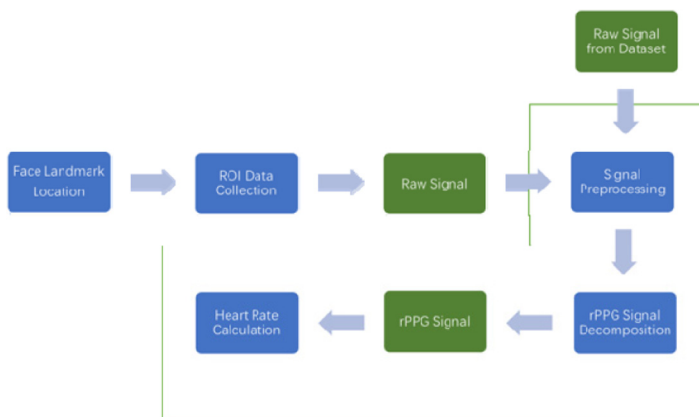


Figure 7.
Evaluation of signal
processing accuracy

Figure 8.
Clean PPG signal and
generated signal with
Gaussian noise

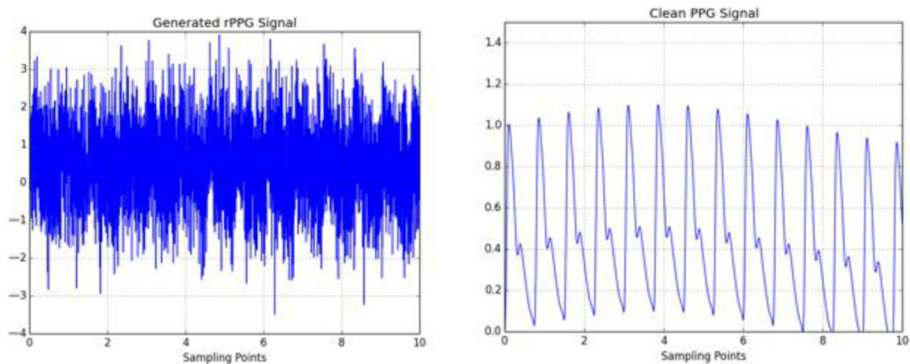
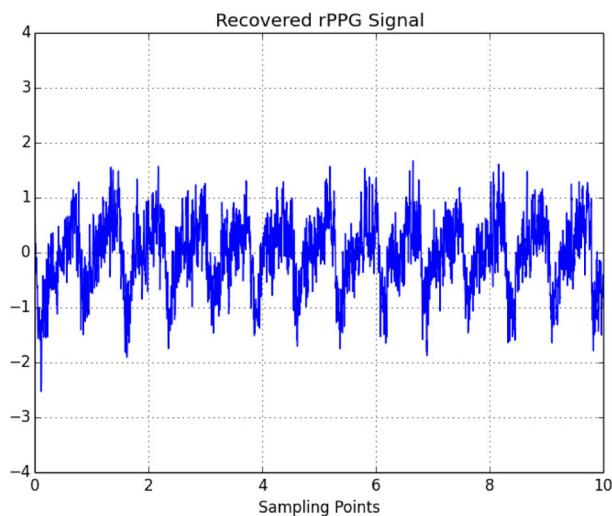


Figure 9.
Recovered rPPG
signal



the calculated the HR from artificially contaminated signals. The comparison between experimental result and ground truth is in [Figure 10](#).

Some of the results are listed in [Table II](#) to show the deviation level.

The evaluation on synthetic data set can strongly prove that the signal processing part of our framework has very good performance and can successfully separate the target signal with its key information and our framework can work under severe sensor noises and successfully calculate the accurate HR.

5.4 Comparison with related works

Due to the lack of source codes and the difference of application usage, we hereby compare the proposed framework with other related works on the qualitative level. Compared to the previous HR measurement method, this proposed framework is accurate with a relative error within 1. And its robustness enables it to be applied for home nursing assistance which is hardly possible for other methods. The rapid

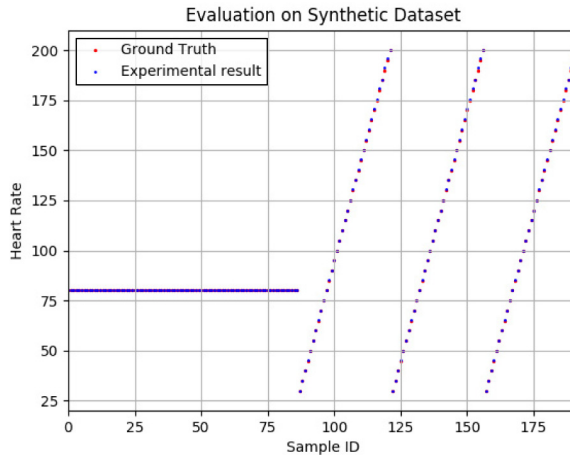


Figure 10.
Output result and
ground truth
comparison

Beats per minute	Experimental result	Ground truth	Deviation
Object 1	40	40.00076	0.00076
Object 2	50	50.00095	0.00095
Object 3	60	60.00114	0.00114
Object 4	70	70.00133	0.00133
Object 5	80	80.00152	0.00152
Object 6	100	100.0019	0.0019
Object 7	120	120.0023	0.0023
Object 8	140	140.2884	0.2884
Object 9	160	160.2888	0.2888
Object 10	180	180.8606	0.8606

Table II.
Data set evaluation
result

measurement speed which can output first result within 5 s also outperforms other works and greatly improve the user experience.

6. Conclusion

In conclusion, a non-contact real-time framework designed for home nursing assistant robots is proposed and validated to be efficient. The framework can detect human HR from a distance under various circumstances including during daily conversation and is robust even with body swing and head motions allowing the users to be at ease. The HR value can be calculated in real-time and a heartbeat curve can be produced at the same time. A low-cost but efficient BSS method is applied in our framework. We evaluated our framework in real-world scenario, inviting dozens of volunteers to take part in the evaluation and successfully proved its robustness and accuracy. The framework has also been validated on data set to verify the correctness of signal processing part. In all, the contribution of this paper is proposing a non-contact HR framework which is suitable for home nursing assistant robots.

However, there are still some limitations that should be taken into serious consideration. First, motion artifacts remain to be a heavy contamination, although we skillfully avoided to

be entangled in such problem. Second, the elaborate information of rPPG signal including the amplitudes is very useful when analyzing other physiological indexed such as oxygen saturation (SpO₂). Third, most of the volunteers are health young and middle-aged people. Few of them have potential cardiac problems so the effect of early warning the physical discomfort remains unknown. In the future, we will firstly supplement the evaluation of the framework. We will also devote more effort into better signal decomposition methods and recover better rPPG signal. This framework will be open sourced soon to contribute to the community.

References

- All About Heart Rate (Pulse). American Heart Association. 22 Aug 2017. Retrieved 25 Jan (2018)
- Charlton, P.H., Bonnici, T., Tarassenko, L., Clifton, D.A., Beale, R. and Watkinson, P.J. (2016), "An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram", *Physiological Measurement*, Vol. 37 No. 4, pp. 610-626.
- Dwyer, R. (1984), "Use of the kurtosis statistic in the frequency domain as an aid in detecting random signals", in *IEEE Journal of Oceanic Engineering*, Vol. 9 No. 2, pp. 85-92, doi: 10.1109/JOE.1984.1145602.
- Elgendi, M. (2012), "On the analysis of fingertip photoplethysmogram signals", *Current Cardiology Reviews*, Vol. 8 No. 1, pp. 14-25, PMC. Web. 26 Feb. 2018.
- Jiang, W.J., Gao, S.C., Wittek, P. and Zhao, L. (2014), "Real-time quantifying heart beat rate from facial video recording on a smart phone using kalman filters", *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom), Natal*, pp. 393-396.
- Kessler, V., Kächele, M., Meudt, S., Schwenker, F. and Palm, G. (2016), "Machine learning driven heart rate detection with camera photoplethysmography in time domain", *Artificial Neural Networks in Pattern Recognition*, Springer, Cham, pp. 324-334.
- King, D.E. (2009), "Dlib-ml: a machine learning toolkit", *Journal of Machine Learning Research*, Vol. 10 No. jul, pp. 1755-1758.
- Kwon, S., Kim, H. and Park, K.S. (2012), "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone", *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA*, pp. 2174-2177.
- Lee, C.M. and Zhang, Y.T. (2003), "Reduction of motion artifacts from photoplethysmographic recordings using a wavelet denoising approach", *IEEE EMBS Asian-Pacific Conference on Biomedical Engineering, 2003*, pp. 194-195.
- Poh, M.Z., McDuff, D.J. and Picard, R.W. (2011), "Advancements in noncontact, multiparameter physiological measurements using a webcam", *IEEE Transactions on Biomedical Engineering*, Vol. 58 No. 1, pp. 7-11.
- Qi, H., Guo, Z., Chen, X., Shen, Z. and Wang, Z.J. (2017), "Video-based human heart rate measurement using joint blind source separation", *Biomedical Signal Processing and Control*, Vol. 31, pp. 309-320.
- Viola, P. and Jones, M. (2001), "Rapid object detection using a boosted Cascade of simple features", *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, pp.1-511-1-518 Vol. 1.
- Wei, C., Woo, W.L. and Dlay, S.S. (2007), "Nonlinear underdetermined blind signal separation using bayesian neural network approach", *Digital Signal Processing*, Vol. 17 No. 1, pp. 50-68.
- Wu, H., Wang, T., Dai, T., Lin, Y. and Wang, Y. (2018), "A Real-Time Vision-Based heart rate measurement framework for home nursing assistance", *To Appear in ICAA*, Vol. 2018.

Further reading

Wikipedia contributors. "Blind signal separation." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 23 Jan. 2018. Web. 21 Feb (2018),

Wikipedia contributors. "Butterworth filter." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 5 Jan. 2018. Web. 22 Feb (2018),

Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology Circulation (1996), 93:1043-1065, originally published March 1, 1996.

About the author

The authors are from School of Electronic Engineering and Computer Science, Peking University, China. The corresponding author of this paper is Associate Professor Tao Wang. Tao Wang can be contacted at: wangtao@pku.edu.cn

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com