A Noncontact Measurement of Cardiac Pulse Based on PhotoPlethysmoGraphy

Xiaohua Wu, Xin Li(✉), Yulin Xu, and Lang Zhang

School of Mechatronics Engineering and Automation, Shanghai University, Shanghai 200072, China
su_xinli@aliyun.com

Abstract. Heart rate measurement is important for monitoring people's physiological and body state. In this paper, a heart rate measurement methodology based on PhotoPlethysmoGraphy (PPG) signal is proposed. Human face positions are detected and tracked in real time by using facial color videos taken from cameras by non-contact shooting. Signals containing pulse components are extracted from images of the forehead skin area for the purpose of calculating blood volume pulse waves via wavelet filtering. Hence, heart rates are calculated after energy spectrum analysis using Fourier transform. The method realizes non-contact measurement, which avoids potential discomfort caused by direct skin contact, and has the advantages of simple operation and low costs. The result indicates that it is sensible to apply this method to daily family heart rate monitoring and remote medical monitoring equipment.

Keywords: PPG · Face detection and tracking · Heart rate · Noncontact

1 Introduction

Family medical services in today more and more disciplines more and more. The remote monitoring of vital signs includes not only high-precision diagnostic equipment, but also simple diagnostic equipment, and for everyone to facilitate. One of the most common examinations in health care monitoring is cardiac pulse measurement. There are many different ways to measure heart rate contact measurements, where the gold standard is ECG. However, the heart of records generated by the potential need for proper application of the electrodes, which may be too complicated and annoying in a home environment. Other methods of measuring cardiac pulse include thermal imaging [1], Doppler optical [2] and ultrasound [3] or piezoelectric measurements. However, these methods are too expensive and can not be used for home care, so finding low-cost, non-contact heart rate measurement is the key to solving the problem.

Photoplethysmography is a new type of photoelectric detection method, first proposed by Schmitt et al. in 2000, and applied to the detection of skin blood flow and related fluctuations in the phenomenon, to study the skin surface wound healing

This work is supported by science and technology commission of shanghai (15411953500).

M. Fei et al. (Eds.): LSMS/ICSEE 2017, Part I, CCIS 761, pp. 13–22, 2017.
DOI: 10.1007/978-981-10-6370-1_2

Happening. According to Poh et al. [4], Harvard and MIT Joint Research Group, 2010, a method of measuring the heart rate and respiration rate based on independent component analysis (ICA) was proposed. The motion artifacts caused by motion disturbances seriously affect the accuracy of the measurement results [5]. The In this paper, an improved measurement algorithm is proposed, and the accuracy of the experiment is evaluated by experiment.

The paper is organized as follows: Sect. 2 talks about the related work for the topics discussed in this paper. Section 3 describes the results of the experiments done and compared with the ground truth. Section 4 provides a conclusion and outlook on future works.

2 Noncontact Cardiac Pulse Measurement

2.1 Photoplethysmography (PPG)

Photoplethysmography (PPG) is an optical measurement technique that can be used to detect blood volume changes in the micro vascular bed of tissue. It has widespread clinical application, for example in pulse oximeters, vascular diagnostics and digital beat-to-beat blood pressure measurement systems. The pulsatile component of the PPG waveform is often called the "AC" component and usually has its fundamental frequency, typically around 1 Hz, depending on heart rate (Fig. 1). This AC component is superimposed onto a large quasi-DC component that relates to the tissues and to the average blood volume. This DC component varies slowly due to respiration, vasomotor activity and vasoconstrictor waves. With suitable electronic filtering and amplification both the AC and DC can be extracted for subsequent pulse wave analysis.

Fig. 1. The pulsatile (AC) component of the PPG signal and corresponding electrocardiogram (ECG). The AC component is actually superimposed on a much larger quasi-DC component that relates to the tissues and to the average blood volume within the sample. It represents the increased light attenuation associated with the increase in microvascular blood volume with each heartbeat.

2.2 Measurement Procedures

An ordinary camera is used to capture face image information in real time with the help of the face detection algorithm designed to detect the human face. Besides, the human face is tracked via the tracking algorithm. After the acquisition of signals from the forehead area of the face image, the ROI image of each frame is divided into three primary

colors to generate *R*, *G* and *B* three channel images. Take the gray space average value of each channel as the signal value of the frame image. Consequently, here come the three original signals *R (t)*, *G (t)* and *B (t)*, which will change into three independent source signals after blind signal separation. Then the correlation analysis of the three independent source signals is carried out, and the signal containing the maximum blood volume change is selected as the signal to be analyzed. Lastly, the signal is analyzed via frequency analysis to obtain the heart rate.

Measurement will only need household PCs and ordinary consumer class cameras with subjects to be located in front of the camera at *0.5* m, facing the camera without significant movements. The measurement devices are shown in Fig. 2.

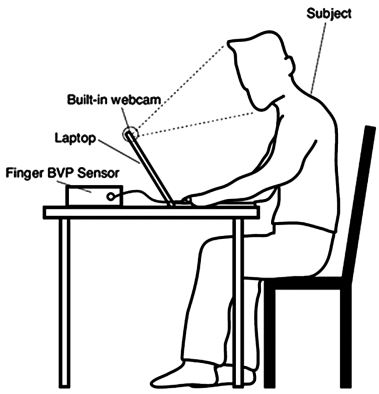

Fig. 2. Experimental setup

First, set the camera resolution pixel as 640×480, frame rate of *30* frames/s and the image color space as RGB. After getting *450* frames of face video from shooting *15* s, process the face video, extracting source signals which contain pulse components. Then wavelet filter for blood volume pulse waves. Finally, analyze heart rate by energy spectrum analysis. The algorithm flow chart is shown in Fig. 3, which mainly includes the following parts: face detection and tracking in Images; choosing the Region of Interest; source signal extraction; signal filtering based on Wavelet Transform; Heart rate calculation based on energy spectrum analysis.

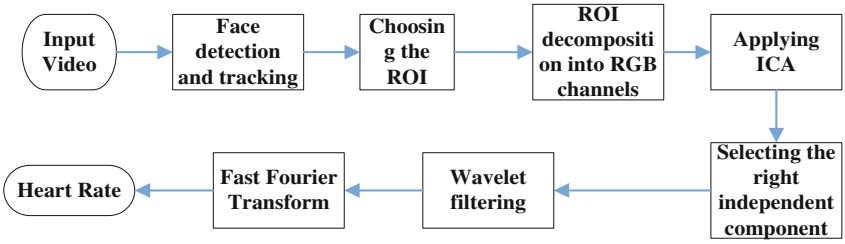

Fig. 3. The algorithm flowchart

2.3 Face Detection and Tracking

One of the main challenges in video-based non-contact heart rate measurement systems is the presence of motion artifacts [8], which has great influence on the accuracy of the whole system and cannot be completely eliminated. In the process of collecting videos, even if the subjects are asked to remain still, some tiny jitter is unavoidable, which will produce great interference on the extraction of weak PPG signals. Therefore, this paper selects the Adaboost algorithm to detect faces, and then combines the KLT feature point tracking algorithm to experiment with face tracking.

The specific algorithm is as follows: Capture video frames, to shorten the processing time to ensure tracking effect, detect human face every *300* frames based on Adaboost to return the location of the face in the image, positioning target tracking area; KLT feature points are used to track the feature points of the target area to achieve face tracking and the rest of the video frame only needs to locate the target tracking region in the next image according to the feature points of the target tracking region in the video image. In order to avoid the loss of tracking, the feature points are updated every *300* frames, and the number of tracking points is set as less than 4 thus the Adaboost tracking algorithm is used to update the feature points of the target tracking region. This iteration to achieve face tracking can effectively shorten the calculation time. The Specific face detection and tracking algorithm flow chart is shown in Fig. 4.

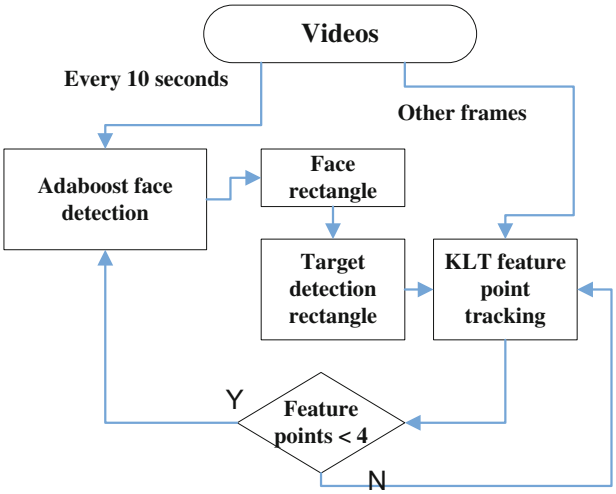

Fig. 4. Face detection and tracking algorithm flow chart

Suppose the coordinate of face detection in video images is *(x, y, w, h)*, wherein *x* represents the face in the image horizontal coordinate position, *y* the face position in the image vertical, *w* face frame width and *h* face frame height. According to face distribution proportion, adjust the target area coordinates to *(x', y', w', h')*, which is shown in Fig. 5(a) while the relation between *(x', y', w', h')* position and face tracking *(x, y, w, h)*

is $x' = x + 0.2 \times w$, $y' = y + 0.25 \times h$, $w' = 0.58$ w, $h' = 0.58$ h. The rectangular box calibration area shown in Fig. 5(b) is the new target tracking area.

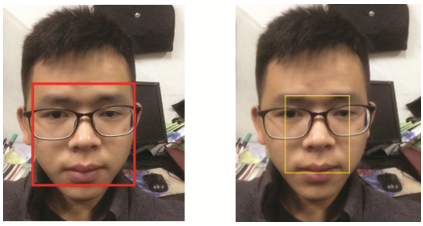

Fig. 5. (a) Face tracking area (b) New target tracking area

In the new target tracking area, the region with rich features such as the eyebrows, mouth and nose is selected which can exclude backgrounds, face edges, hairs and other tracking feature points of interference. Therefore, it can avoid the late ROI tracking deviation from the face area, to achieve a better ROI tracking and positioning, thus inhibit the movement of artifacts.

2.4 Choosing the Region of Interest (ROI)

Previous HR measurement methods use the Viola-Jones face detector [9] of Opencv [10] to detect faces. It only finds coarse face locations as rectangles, which is not precise enough for HR measurement task since non-face pixels at corners of rectangles are always included. The case becomes even worse when the face rotates. To this end, we first apply Adaboost face detector to detect the face rectangle on the first frame of the input video, Then select a rectangle as the region of interest in the forehead according to the method shown in Fig. 6(b).

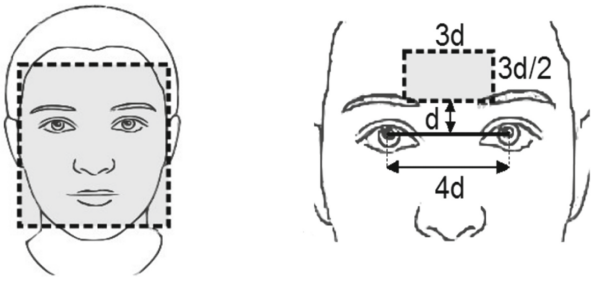

Fig. 6. Definition of ROI's for two approaches utilized in the paper, (a) the whole face as a ROI, (b) selected part of the forehead as a ROI

Two rules are followed for defining the ROI: the first one is to exclude the eye region since blinking may interfere with the estimated HR frequency; the second one is to indent

the ROI boundary from the face boundary, otherwise non-face pixels from background might be included during the tracking process.

2.5 Recovery of BVP from Video Recording

In the current commonly used method [6], the source signal is separated by color separation of the image, and a color channel was selected for spatial pixel averaging. ICA [7] is a technique used for isolating independent signals from a set of vectors that consist of a linear combination of these signals [11, 12]. The difference between ICA and other methods is that ICA searches for statistically independent and non-Gaussian signals [13]. ICA is applied to many applications in various areas.

When taking a facial video, RGB (red, green and blue) sensors will capture a blend of the reflected PPG signal along with other noise originating artifacts. Each color sensor will record a blend of the initial source signals with a little change in their weight as a result of the differences in the hemoglobin absorptivity in the visible and near-infrared spectral range. In our method, the detected signals from the RGB sensors are denoted as $y_1(t)$, $y_2(t)$ and $y_3(t)$ which represent the amplitudes of the saved signals at time t. Also, it supposes that $x_1(t)$, $x_2(t)$ and $x_3(t)$ are the three fundamental source signals that have been linearly combined to generate $y_1(t)$, $y_2(t)$ and $y_3(t)$. Hence, Eq. (1) describes the aforementioned relationship between captured and source signals.

$$\mathrm{y(t)} = \mathbf{A}\mathrm{x(t)} \tag{1}$$

Where $y(t)$ is vector $[y_1(t), y_2(t), y_3(t)]^T$, $x(t)$ is vector $[x_1(t), x_2(t), x_3(t)]^T$ and $\boldsymbol{A}$ is a square matrix (called mixture matrix) of size 3×3 that include the mixture coefficients a_{ij}. Therefore, in order to model the relationship between $y(t)$ and $x(t)$, we need to determine a separation matrix $\boldsymbol{W}$ that is estimates the inverse of the mixture matrix $\boldsymbol{A}$. Once $\boldsymbol{W}$ is determined, the source signals (i.e. $x(t)$) can be estimated. Some of the independent random variables is more Gaussian comparing to the original variables, that is why the non-Gaussian of all sources should be maximized by $\boldsymbol{W}$ to reveal the independent sources (Fig. 7).

The three independent source signals obtained from ICA are disordered. Since the green signal best reflects heart beats, the independent source signal is correlated with the green channel signal and the most relevant independent signal is selected as the effective signal U(t).

2.6 Signal Filtering Based on Wavelet Transform

In order to filter the baseline drift and high frequency noise in the source signal, wavelet transform and wavelet inverse transform are used to filter the passband frequency of 0.75–2.0 Hz (corresponding to heart rate of 45–120 times/min).

Wavelet transform (WT), a time-frequency analysis method, is similar to Fourier transform. However, unlike the Fourier transform, the CWT can detect the abrupt changes in the signal due to the variable sampling step, so it is more suitable for the analysis of biomedical signals. Wavelet transform can be divided into Continuous

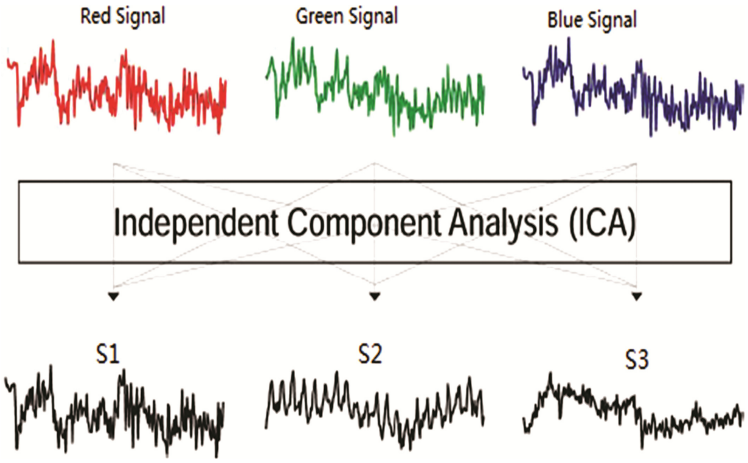

Fig. 7. After the raw signals are detrended and normalized, ICA is applied to separate three independent sources.

Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). Since the DWT discretizes the scale factor and the translation factor in the CWT, some frequency information may be ignored. Therefore, the system chooses CWT to convolution the source signal *u(t)* and $\psi_{\tau,s}$ according to the Eq. (2). Where $\psi_{\tau,s}$ is a wavelet basis function, derived from the mother wavelet translation, *s* is the scale factor that is the expansion factor, and τ is the translation factor. The scale factor is related to the frequency, the larger the scale factor is, the longer the sampling step is whereas the lower the frequency resolution is. At present, the commonly used wavelet basis functions include *sym* wavelet, *dB* wavelet, *Haar* wavelet, *Morlet* wavelet and so on. The selection of wavelet basis function should be considered from two aspects: general principles and specific objects. The general principles are: (1) orthogonality; (2) compact support; (3) symmetry; (4) smoothness. But it is very difficult to fully satisfy these principles, as compact support contradicts with smoothness and the compactness of orthogonality makes it impossible to achieve symmetry, so it is reasonable and meaningful to find a way to properly balance these principles. Choices are made depending on processing signals. *Morlet* wavelet is validated for PPG signal analysis [20].

$$\begin{cases} CWT_G^{\psi}(\tau, s) = \int_{-\infty}^{+\infty} u(t)\psi_{\tau,s}(t)dt \\ \psi_{\tau,s}(t) = \dfrac{1}{\sqrt{|s|}}\psi\left(\dfrac{t-\tau}{s}\right) \end{cases} \tag{2}$$

In the Eq. (2), the inverse wavelet transform is used to inverse transform the signal in the scale range of 0.75–2 Hz, and the reconstructed signal is the pulse wave BVP, as shown in Fig. 8.

$$\begin{cases} \mathrm{BVP} = \dfrac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} \mathrm{CWT}_G^{\psi}(\tau, s) \dfrac{1}{\sqrt{|s|}} \psi\left(\dfrac{t-\tau}{s}\right) d\tau ds \\ C\psi = \int_0^{\infty} \dfrac{|\hat{\psi}(\zeta)|^2}{|\zeta|} d\zeta < \infty \end{cases} \tag{3}$$

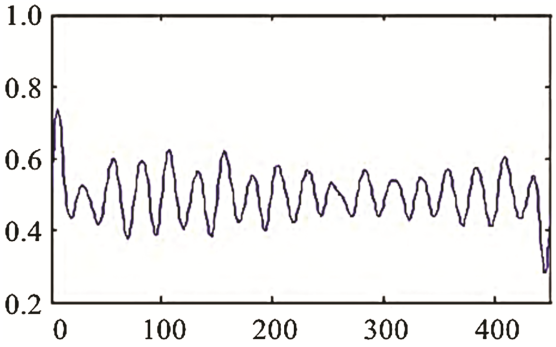

Fig. 8. BVP wave

2.7 Heart Rate Calculation Based on Energy Spectrum Analysis

The frequency domain is the main interest; therefore, to convert the time domain signal to the frequency, the Discrete Fourier Transform (DFT) is usually used. For less processing time the FFT is applied. The FFT's complexity is much lower than DFT. After performing the FFT, the highest peak detected in the interest band (which is the HR band between *45* and *120*). The detected peak is then converted to the right frequency in the FFT vector. In the Fig. 9, $HR = 1.203 \times 60 \approx 72$ beats/min.

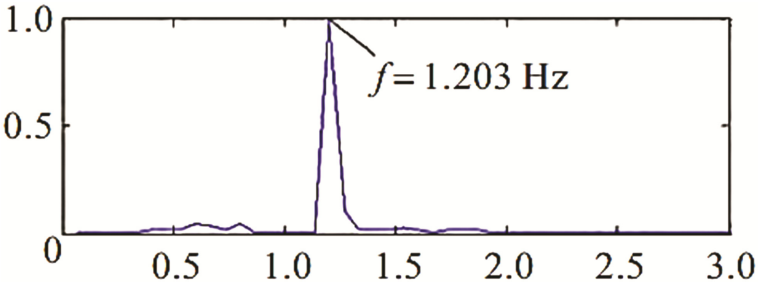

Fig. 9. Energy spectrum of BVP

3 Experiments

In this experiment, *20* subjects (*15* males and *5* females) were recruited. The selected subjects are between *20* and *30* years old. They were tested in two groups under the uniform distribution of light. In the first group, the *20* subjects' heart rates were detected in a resting state, and the subjects were measured by the finger BVP sensor at the same time. In the second group, the *20* subjects' heart rates were detected after *50* deep squats, and the subjects were measured by the finger BVP sensor at the same time. The data obtained are shown in Table 1.

Table 1. Accuracy experiment

The first group				The second group			
Sub. (Resting state)	Our method	Finger BVP sensor	Error	Sub. (After exercise)	Our method	Finger BVP sensor	Error
1	66	63	3	1	113	116	3
2	54	59	5	2	94	99	5
3	81	77	4	3	110	108	2
4	53	52	1	4	100	98	2
5	50	52	2	5	95	99	4
6	74	74	0	6	104	100	4
7	60	61	1	7	101	100	1
8	71	70	1	8	110	113	3
9	56	58	2	9	100	99	1
10	58	58	0	10	112	110	2
11	54	55	1	11	90	93	3
12	50	53	3	12	99	98	1
13	62	62	0	13	103	102	1
14	68	67	1	14	105	105	0
15	59	62	2	15	98	97	1
16	72	71	1	16	103	100	3
17	65	66	1	17	92	97	5
18	80	75	5	18	105	103	2
19	71	69	2	19	101	102	1
20	63	63	0	20	99	98	1

As can be seen from Table 1, the maximum error of the measured value of this method is about 5 beats/min; the average error is about 2 beats/min. The average correlation coefficient between the measurement results by our method and by the ECG measurements was 0.982, and the result meets the requirements of the People's Republic of China's pharmaceutical industry standards (error $\leq$ 5 beats/min). So this method can be used in home medical system.

4 Conclusions

This method intends to achieve accurate non-contact heart rate measurement with the help of ordinary color camera under natural light conditions. It is a good choice when traditional measurement methods are not suited to some subjects. Besides, in remote medical inquiry system, this method can monitor heart rates whenever necessary. Hence, it has a bright prospect in the future family medical service. However, in the case of too short video length and poor lighting conditions, this method may not measure heart rates accurately. Therefore, in order to improve the robustness of heart rate measurement, further study is needed to solve these problems.

References

1. Garbey, M., et al.: Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. IEEE Trans. Biomed. Eng. **54**(8), 1418–1426 (2007)
2. Shastri, D., et al.: Imaging facial signs of neurophysiological responses. IEEE Trans. Biomed. Eng. **56**(2), 477–484 (2009)
3. Holdsworth, D.W., et al.: Characterization of common carotid artery blood-flow waveforms in normal human subjects. Physiol. Meas. **20**(3), 219 (1999)
4. Poh, M.Z., Mcduff, D.J., Picard, R.W.: Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Trans. Biomed. Eng. **58**(1), 7 (2011)
5. Lewandowska, M., et al.: Measuring pulse rate with a webcam - a non-contact method for evaluating cardiac activity. In: Proceedings Federated Conference on Computer Science and Information Systems - FedCSIS 2011, Szczecin, Poland, 18–21 September 2011, DBLP, pp. 405–410 (2011)
6. Tsouri, G.R., et al.: Constrained independent component analysis approach to nonobtrusive pulse rate measurements. J. Biomed. Optics **17**(7), 077011 (2012)
7. Allen, J.: Photoplethysmography and its application in clinical physiological measurement. Physiol. Meas. **28**(3), R1–R39 (2007)
8. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, vol. 1, pp. I-511–I-518. IEEE (2001)
9. Bradski, G.: The Opencv library. Doct. Dobbs J. **25**(11), 384–386 (2000)
10. Addison, P.S., Watson, J.N.: Rapid Communication: a novel time frequency-based 3D Lissajous figure method and its application to the determination of oxygen saturation from the photoplethysmogram. Meas. Sci. Technol. **15**(11), L15 (2004)
11. Comon, P.: Independent component analysis, a new concept? Sig. Process. **36**(3), 287–314 (1994)
12. Stone, J.V.: Independent component analysis. Trends Cogn. Sci. **6**(1), 529 (2005)
13. Kwon, S., Kim, H., Park, K.S.: Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. In: International Conference of the IEEE Engineering in Medicine & Biology Society PubMed, pp. 2174–2177 (2012)

http://www.springer.com/978-981-10-6369-5

Advanced Computational Methods in Life System Modeling and Simulation
International Conference on Life System Modeling and Simulation, LSMS 2017 and International Conference on Intelligent Computing for Sustainable Energy and Environment, ICSEE 2017, Nanjing, China, September 22-24, 2017, Proceedings, Part I
Fei, M.; Ma, S.; Li, X.; Sun, X.; Jia, L.; Su, Z. (Eds.)
2017, XVIII, 609 p. 348 illus., Softcover
ISBN: 978-981-10-6369-5