Heart Rate Measurement

Parth Kansara, Ritwik Dhar, Riddhi Shah, Devansh Mehta, Purva Raut
Dwarkadas J. Sanghvi College of Engineering, Mumbai - 400101, India
E-mail: parthskansara@gmail.com, ritwik1798@gmail.com, riddhishah2103@gmail.com, devanshmehta118@gmail.com, purvapraut@gmail.com

Abstract. In recent studies on non-invasive techniques for heart rate measurements, various Computer Vision algorithms based on Ballistocardiography (BCG) have been employed. This method captures minimal head motions from facial videos, that result from the pumping of blood to the head through the carotid arteries, at each cardiac cycle. We move towards BCG because the conventional technique of Photoplethysmography (PPG) fails to yield accurate results from facial video in case of skin color variations. This paper proposes an improved system for accurately measuring the heart rate and heart rate variability to infer important information about the subject’s health. It incorporates functions from the Dlib toolkit, which provide robust face detection along with facial landmark tracking. Relevant data in terms of facial video and ground truth was acquired from 5 test subjects, in 3 states - sitting, standing and post exercise. The system exhibits promising results when validated using a wearable smart watch with inbuilt heart rate sensor.

1. Introduction
Estimation of the heart rate is an important component for assessing a person / individual’s physiological and pathological state. The heart function is to pump blood around the body by beating at about 60 to 100 beats per minute (bpm). Heart rate greater than 100 bpm (Tachycardia) and heart rate lower than 60 bpm (Bradydcardia) are considered as irregularities. Cardiac abnormalities are caused by physiological and pathological conditions that alter the normal electrical impulse that controls the pumping action of the heart. Tachycardia is typically caused by elevated blood pressure, cigarettes, fatigue, acute stress, drug side effects and heart disease damage to heart tissues. Bradycardia is typically a source of ageing-related damage to hypertension, rheumatic fever, obstructive sleep apnea, myocarditis and cardiac tissue. To any of these conditions the risk factor will result fatally.

Since remote sensing has evolved, the research on non-contact heart rate calculation has attained more emphasis. Remote sensing has reduced the amount of wiring involved with the heart rate monitoring methods dependent on interaction. Because of its reasonable price and non-invasive design for calculating the heart rate, heart rate estimation using a digital camera sensor is a rapidly growing research field. The measurement of heart rate using digital camera sensor is mainly based on the photoplethysmography and ballistocardiography principles. Digital camera-based heart rate measurement techniques have developed with the use of mobile phone and webcam sensors with high resolution charging coupling system (CCD) and complementary metal-oxide semiconductors (CMOS) sensors. Hence, calculating heart rate from facial images using digital camera sensors has been one of the inspiring avenues in a non-invasive method for
collecting physiological signals.

Ballistocardiography (BCG) is a non-invasive procedure for producing a graphical image of the repetitive movements of the human body caused by heartbeat. These repetitive movements occur due to the accelerated flow of blood as it is pumped and transferred through the body’s great vessels during cycles of relaxation and contraction, respectively called diastole and systole. In other words, BCG may provide details on the overall efficiency of the circulatory system; this is because BCG tests the fluid flow, i.e. the density of the flowing blood and the heart during the cardiac cycle. During atrial systole, the body’s center of mass travels towards the head of the body as the blood is pumped into the large vessels. In other cases, as the blood flows into the peripheral arteries and focuses with in peripheral veins far away from the heart, the center of mass shifts toward the foot. As a function of heart operation, oxygen levels, and body movements, this change involves many factors. This transfer of the body’s center of mass produces the BCG waveform as the movement of blood varies during the cardiac cycle.

Due to the revolution in information technology, BCG has at present been granted a lot of attention, covering hardware technologies as well as applications and services. BCG devices like electromechanical film sensors, load cells, piezoelectric polyvinylidene fluoride sensors, hydraulic sensors, pneumatic sensors, and fiber optic sensors can be installed in environmental conditions without the need for the intervention of the medical personnel. Therefore, the effect on emerging e-health programmes is excellent. Ultimately, BCG helps alleviate the burden of checkups and the patient’s feelings and reactions to stimuli. To sum up, BCG can be very useful in many applications such as controlling heart activity and efficiency, in addition to monitoring breathing disorders by sleep. One of BCG’s most popular functionality is simplicity and ready connectivity, which enables users to install the device in their homes without compromising the confidentiality and everyday activities of the people.

The advancement of calculating heart rate using facial images has provided a new approach to fitness screening in the areas of health treatment, telemedicine, recovery, exercise and ergonomics. This paper discusses the evolution and showcases the state of the art on heart rate measurement via facial videos. The paper has been divided into six sections. Section 2 gives in-depth overview of Ballistocardiography with emphasis on conventional methodologies followed and related work in the field. Section 3 focuses on our model implementation, describing the system architecture and signal processing algorithms. Section 4 gives the system performance on various test subjects in multiple conditions and error rates. Section 5 shows extraction of heart rate variability from the data previously gathered and its significance. The paper concludes with the inference and future additions to be implemented to improve the research in future.

2. BCG

2.1. Conventional methods

In particular, the performance of the above-mentioned BCG sensors is a composite signal consisting of heart activity, breathing activity, and body movements. Those three impulses should then be isolated from each other in order to calculate vital signs. The separation method is typically achieved by adding a band-pass filter with varying cutoff frequencies as per the signal range of interest. In other instances, wavelet multiresolution analysis or decomposition algorithms like empirical mode decomposition can be used to achieve the separation process. It should be known that during body movements critical events cannot be observed, and so they should be omitted before the measuring process. Going with the separation process, i.e. the gathering of heart signals and respiratory signals, multiple algorithms for vital measurements can then be applied.

They are usually divided into the following categories: algorithms in the time-domain, algorithms in the frequency-domain and algorithms in the wavelet domain. Below is a description of these algorithms to illustrate which type of algorithms to be used in our analysis. Firstly,
time-domain algorithms are predominantly based on observing local maxima or local minima using a moving frame, and thus discovering the interval between the dominant BCG signal J-peaks. Conversely, due to the nonlinear and non-stationary nature of the BCG signal this method has certain drawbacks. The assumption is that the BCG signal does not show stable J-peaks, as would generally be the case for in-home monitor overnight. Additionally, activity artefacts would certainly impact its precision.

Secondly, frequency-domain algorithms do not provide details on interbeat periods. They may also provide details about fluctuations in heart rate. Typically, this is achieved by taking the fast Fourier transform (FFT) of the approximate frequency logarithm, i.e. cepstrum of the signal using a sliding window. The dominant frequency in a given frequency spectrum is then obtained. The disadvantage of these equations would be that the peak in the continuum can become larger and several peaks will occur which can create a difficulty in the calculation of the vital signs.

Finally, the purpose of wavelet-domain algorithms is just to decompose the signal into different components. Therefore, it is possible to pick the component that displays an agreement with the vital signs. In other words, the part chosen provides only details on the cardiac cycles or respiratory cycles, respectively. Interbeat intervals can be conveniently detected by merely using a peak detector. Also effectively used for electrocardiogram signal processing was wavelet analysis along with other techniques such as denoising, compression, and biometric recognition. An alternative approach to wavelet decomposition is empirical decomposition, which is also a very effective way to dealing with non-stationary signals like cardiorespiratory signals. In addition to the aforementioned algorithms, machine learning methods for calculating heartbeats have been applied. However, a restricting property is the manual marking of training results. Also, the training phase should be repeated if the data collecting process has modified.

2.2. Conventional methods
Balakrishnan et al., suggested a system for estimating heart rate using BCG-based facial imagery in 2013. This approach focussed on capturing in the video the unwanted head move. The heart rate from the facial video was collected by measuring feature point velocities on the area around the face. Extraction of the velocity of the function points to map the head’s microscopic motion. The head feature points were derived from the authors specified region of interest (ROI). The writers used a mixture of regions above the line of the eye (i.e. forehead) and the area below the line of the eye that included the cheeks and the upper portion of the lips. The reason for ROI selection is that many of the capillaries which branch out of the carotid arteries were located in the selected region. The eye line was omitted because the movement of the eye in the derived BCG signal would result in motion artefacts.
The writers used the face detector Viola-Jones, to identify the mask and collect the ROI. The ROI feature point was tracked with the help of face tracking algorithm Kanade Lucas Tomasi (KLT). Temporarily filtered the tracked points using a bandpass philtre which worked at 0.75 Hz to 2 Hz. The operating frequency was selected to match cardiac frequencies between 45 to 120 bpm. The primary study of the components (PCA) was used to remove artefacts and to collect the BCG as a periodic signal. Consequently, a simple peak detection algorithm had calculated the subject’s heart rate.

Shan et al. suggested a BCG method of extraction on the basis of Balarishnan et al. The writers used a single ROI here, instead of a ROI mix. The front region was mainly used to remove the feature points because the area below the eye line was much more vulnerable to motion errors, induced through non-rigid facial muscle movements related to speech or emotional gestures. In addition, the BCG signal was derived using independent component analysis (ICA) instead of PCA. Lastly, heart rate is calculated by removing the BCG signal from the power spectral density (PSD). While there is not much variation between the findings obtained using the PCA or ICA approaches, the major benefit of the BCG based technique was that the process itself was invariant to the variance in lighting and skin tone variations.

By using deformable face fitting algorithm, Haque et al. suggested a BCG based heart rate prediction method using fusing corner feature points with good features to map and predefined facial landmarks. The writers used the feature-based fusion technique to address the drawbacks with respect to the lack of features due to facial shifting or occlusion. Tracking the function points using the KLT monitoring algorithm and removing the vertical components of the trajectories as raw traces of BCG. The authors further improved the raw traces by using a bandpass filtering method at operating frequencies of 0.75 Hz to 5 Hz, using an eight-order Butterworth band-pass filter. Further optimizing the filtered BCG signals by adding a moving average filter to eliminate artefacts of motion. The writers used a moving average filter with a window size of 300 frames / samples. The authors applied PCA from raw BCG signals to approximate the linearly uncorrelated structure.

The heart rate was calculated by using an algorithm suggested by Irani et al. to predict periodicity. The algorithm for periodicity estimation worked by estimating the most frequent PCA portion by conducting a stochastic search using the frequency indexes extracted from the discrete cosine transformation (DCT). The authors, therefore, calculated the heart rate by selecting the first harmonic frequency by applying fast Fourier transform (FFT) to the most frequently used part. BCG-based method of measuring heart rate is also accurate in situations where the mask is not visible. As the BCG-based approach is based primarily on the calculation of the cyclic motion of feature points on the head side. The solution based on BCG is particularly favorable, because the process is invariant to the variation of illumination. The dependency on motion characteristics to approximate the BCG signal, however, brings a significant downside to that strategy. Where the process will be particularly susceptible to modifications in the subject’s voluntary motion. Thus, calculating heart rate during voluntary head movement may reduce the method’s reliability.

3. Implementation
The cardiac-related head motions are minimal and blended in with a number of other involuntary head movements. Involuntary head movement is caused due to multiple factors which disturb the isolation of pulsate activity due to motions.

One is the oscillatory pendulum motion which holds the head in dynamic equilibrium. We find that the vertical orientation is the strongest axis to calculate the pulse-induced shift of the upright head, since the horizontal axis appears to absorb more of the complex balance swaying. The bobbing caused by respiration is a second type of accidental head movement. We cope with this by filtering out activity at low frequencies. The net acceleration of accidental vertical
head movement was estimated at about 10 mG. The average length of one heart cycle’s left ventricular ejection period is around 1/3 seconds. Using those numbers, we can measure an approximation of the rotation of the head to be 5 mm. Although this measure neglects the dynamic head-system structure, it does include an indicator of how minimal the movement is.

3.1. Data acquisition
Our system takes an input video of the user’s head and produces a heart rate as well as a set of beat positions that can be used to measure the frequency of beat to beat. Second, we capture the head movement through feature tracking techniques. The motion corresponding to the pulse is then extracted and mapped onto a 1D signal that helps us to distinguish individual beat limits from trajectory peaks. To do this, we use PCA and pick the part that best fits a pulse with its temporal power spectrum. We project feature points trajectories onto this part and extract the positions of beat as local extreme.

We begin by locating the area of the head and modelling head motion using tracked trajectories of feature points. We use the trajectories vertical part for research. The trajectories have extraneous movements beyond the spectrum of possible pulse speeds at frequencies and so we briefly filter them out. We then use PCA to decompose the patterns into a collection of independent source signals which define the head motion’s main elements. To choose the right source for measuring and estimating the length of individual beats, we analyses the spectra of frequency and select the source with the most distinct major frequency. Using this frequency average pulse rate is defined. We do peak tracking in the time-domain for more fine-grained evaluation and measurement of the beat durations.

3.2. Region Of Interest (ROI)
We noticed several facial feature choices for each subject based on its personal characteristics which could be used to map BCG. Examples of the characteristics we have used consistently
include facial hair, nostril, acne and pigmentation of the skin. The BCGs on the same subject had similarities from various facial features.

Figure 3. Facial Landmarks

We also delete the feature points on the eyes so that blinking artefacts will not impact our performance. Many of the feature points can be unpredictable and contain erroneous trajectories. To maintain the most stable features, we calculate the cumulative distance between consecutive frames traveled at each point and take the average of these feature points as the area of interest to be considered.

Considering the subject’s face is clearly visible in our videos, Dlib’s `get_frontal_face_detector()` function is employed for detecting the face ROI. It uses the `scan_fhog_pyramid()` object, based on fractional Histogram of Oriented Gradients (HOG). After this detection, we call the predictor function from the Dlib toolkit which can accurately predict the facial landmarks that are to be used for tracking.

Figure 4. Comparison of feature tracking algorithms for region of interest

The function predicts 68 landmarks, including corners of eyes, eyebrows, nose, lips and jaws.
Out of these, we do not consider the landmarks associated with the eyes (points 37 to 48 in the figure). This is to eliminate the error introduced by involuntary blinking. Landmarks associated with the lips (points 49 to 68 in the figure) can also be discarded in case the subject is talking or performing any action leading to movement of the lips. The y co-ordinates of the selected 56 landmarks are tracked through each frame between frame 1 and frame \(N\), since that is the primary axis of the Newtonian motion. The 2D array, thus obtained, captures the vertical motion of the 56 landmarks. This is the raw signal that tracks the movement of the head in time.

### 3.3. Signal Processing

Not all trajectory frequencies are needed for pulse detection or are useful. The resting pulse rate for a typical adult fall below \([0.75, 2]\) Hz, or \([45, 120]\) beats / min. We find that frequencies below 0.75 Hz have a negative impact on output of our system. This is because low-frequency motions such as respiration and body changes have a high amplitude and overpower the function points trajectories. Harmonics and other frequencies greater than 2 Hz, however, have valuable accuracy required for peak detection. Taking those elements into account, we filter each \(y_n(t)\) to a \([0.5, 5]\) Hz passband. For its maximally flat passband we use a Butterworth filter of the 5th order.

To separate pulse from the mixture of motions caused by other activities like respiration, vestibular activity and facial expressions, we need to decompose this mixed motion into sub signals. In order to do this, we take the multidimensional head orientation at each frame as a separate data point and use PCA to find a set of key dimensions that differ the orientation along. We then choose a dimension to build the time-series location to receive the pulse signal.

With \(N\) feature points, the \(N\)-dimensional position of the head in motion in the \(t\) coordinate system is expressed as \(G_t = [y_1(t), y_2(t), \ldots, y_N(t)]\). The calculation of mean and covariance matrices are as follows:

\[
\bar{m} = \frac{1}{T} \sum_{i=1}^{T} m_i
\]

\[
\sum_m = \frac{1}{T} \sum_{i=1}^{T} (m_t - \bar{m})(m_t - \bar{m})^T
\]

The eigenvector corresponding to the required ‘\(m\)’ number of largest eigenvalues is contribution rate \(k\) and is considered the principal axis for position change.

\[
k = \frac{\lambda_i}{\sum_{i=1}^{N} \lambda_i}
\]

where \(\lambda_i\) denotes the eigenvalues corresponding to the eigenvectors of the covariance matrix.

The problem persists as to which eigenvector is to be used to collect pulse signal. The eigenvectors from PCA are ordered in such a way that first eigenvector explains the most variation in the results, second most explains, and so on. Although first eigenvector describes much of the variation, we should be able to pick the most periodic signal.

Lomaliza et al. proposed signal selection using a periodicity computation approach. Firstly, fast Fourier transform is applied to the first five principal components \((T_i)\) of the signal. Then, the maximum frequencies \((f_i)\) are used to calculate the maximum period \((T_i)\) as \(T_i = \frac{f_s}{f_i}\), where \(f_s\) is the data sampling rate. A lagged version \(s_i(t)\) for each \(s\) is computed using maximum time period such that:

\[
s_i'(t) = s_i((t + T_i) mod N)
\]

Here, \(N\) denotes the number of data samples (which can be computed as the product of data sampling rate and video length). Ultimately, the periodicity of signal \(s_i\) is calculated by the bivariate correlation with its lagged version \(s_i'\) as shown below:
Here, $T_s$ denotes the data sampling period and $s_i$ and $s_i$ are the sample means. This above shown method indicates that a periodic signal will lag relative to the length of its maximum period.

### 4. Experimental Results

We take into consideration 5 participants (3 male, 2 female) to validate the algorithms applied for heart rate measurement. The model designed is validated under three user conditions - resting stance, standing stance and after exercise stance to assess variability and accuracy of the system. All the testing has been performed under indoor environment as the illumination problem is not much adverse in applications using ballistocardiography (BCG). The webcam resolution is set at 1280x780 pixels with a sampling rate of 30fps. A Xiaomi heart rate smartband was used as a reference system to validate the algorithm performance.

The signal selection approach based on correlation was able to select the correct principal component as the periodicity calculations from Balakrishnan methods was less accurate. The periodicity index does not specifically differ much and low variance often makes it prone to noise. As it can be seen below, the correlation based method provides components more easily differentiable.

![Signal selection algorithm output](image)

**Figure 5.** Signal selection algorithm output

The new method of measuring periodicity has been found to help pick the signal variable compared to the actual heart rate than the approach previously proposed. Thus the mean error was decreased by approximately 2-3 BPM in the new system and approximately 3-4 BPM in the current systems using the modified periodicity calculation process. The proposed method was able to calculate the heart rate with a mean error of less than 0.75 BPM and was reliable enough for use in everyday heart rate tracking. The validation was performed from 5 subjects in the following states:

- Sitting: We have arbitrarily chosen indoor environments for taking videos under nearly uniform lighting.
Standing: For standing recordings, the subject was examined with a laptop webcam at head height in indoor conditions at a specific venue.

Post Exercise: This procedure was carried out after the subject had performed a series of 30-50 skips and then measured in standing posture.

Figure 6. Overall System Setup

To calculate the difference between the observed value and the reference value, root mean square error (RMSE) value is used as shown in the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (f_i - \bar{f}_i)^2}$$  \hspace{1cm} (5)

The relative error is used to stipulate the degree of precision between the calculated values and the reference values as seen in the equation below:

$$\text{error} = \frac{f_{pre} - f_{ref}}{f_{ref}}$$  \hspace{1cm} (6)

Table 1. Comparative results: Sitting

| Participant | Our method | Existing SOA | Reference |
|-------------|------------|--------------|-----------|
| 1 (Male)    | 61.3       | 62.4         | 61        |
| 2 (Male)    | 58.0       | 58.4         | 58        |
| 3 (Female)  | 82.7       | 82.4         | 82        |
| 4 (Male)    | 73.3       | 74.5         | 73        |
| 5 (Female)  | 58.4       | 59.1         | 59        |
Results from the smartwatch, the Balakrishnan method and the suggested method are as seen in Table I. The average values from the proposed method for all subjects are almost equal to the reference rates. The mean error is 1.2 percent, which is less than the Balakrishnan method’s 1.5 percent average error. That indicates that the method proposed is more reliable than the method used by Balakrishnan.

5. Heart Rate Variability

Heart rate variability (HRV) is a commonly used indicator of development in the autonomic nervous system (ANS). The two branches of the ANS are the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS) that dynamically regulate the changes in heart beat-to-beat. The portion of low frequency HRV (LF) is modulated by baroreflex activity and includes both sympathetic and parasympathetic activity. The part of high frequency (HF) represents parasympathetic effect on the heart, it is related to the arrhythmic respiratory sinus (RSA).

To measure HRV, it is important to precisely calculate the interarrival periods of beats, which can be measured by finding the ”R” peaks in successive beats in an electrocardiogram (ECG). A lack of adequate variance while the subject is resting indicates that under stress, the nervous system does not perform well. Patients with diminished HRV have an increased chance of adverse effects such as fatal arrhythmias. The automated identification of heart beat positions is less clear in the case of BCG signals due to the changing anatomy of individual heart beats and increased sensitivity to motion artefacts. Therefore, we used the continuous local interval estimation (CLIE) algorithm to approximate the beat-to-beat intervals and heart beat positions from the BCG signals automatically. This method of measuring heart beats does not rely on fiducial points. Instead beat-to-beat intervals are calculated on a small adaptive processing window (ideally containing two beats) using three separate approaches, such as the auto correlation of the signal. They then integrate these individual figures using a Bayesian approach. The analytics window is moved around the signal using small intervals with respect to

---

Table 2. Comparative results: Standing

| Participant | Our method | Existing SOA | Reference |
|-------------|------------|--------------|-----------|
| 1 (Male)    | 69.8       | 69.5         | 72        |
| 2 (Male)    | 77.1       | 77.8         | 77        |
| 3 (Female)  | 92.2       | 91.7         | 92        |
| 4 (Male)    | 81.6       | 82.2         | 82        |
| 5 (Female)  | 75.3       | 75.5         | 75        |

Table 3. Comparative results: Post Exercise

| Participant | Our method | Existing SOA | Reference |
|-------------|------------|--------------|-----------|
| 1 (Male)    | 119.1      | 118.4        | 120       |
| 2 (Male)    | 142.3      | 139.8        | 141       |
| 3 (Female)  | 133.1      | 135.4        | 136       |
| 4 (Male)    | 115.6      | 117.5        | 117       |
| 5 (Female)  | 129.4      | 130.1        | 129       |
standard interval lengths, resulting in each interval occurring in several consecutive measurement periods.

Objects and outliers have to be omitted before calculating HRV indices from the RR and BCG interval series. BCG beat-to-beat intervals which have been flagged by the CLIE algorithm as objects are then immediately omitted from the HRV computations. Outliers may either be triggered by arrhythmic heart beats or incorrect calculations of intervals in the RR interval series and in the BCG interval series.

The power spectral densities (PSDs) used for measuring the frequency-domain HRV indices were calculated using the periodogram Lomb-Scargle (LS). The LS periodogram is an effective method for evaluating irregular sampled signals, such as data of beat-to-beat intervals, because no interpolation or resampling of the original signal is needed. This property is particularly valuable for the BCG, since BCG-derived beat-to-beat intervals can involve incomplete or defective segments due to motion artefacts on a regular basis. When calculating the PSD using the LS periodogram these parts can be conveniently removed.

6. Summary
In this research, we have built a realistic, head-motion-based (i.e., BCG-based) heart rate monitoring device that can operate on smartphones to overcome the shortcomings of previous heart rate monitoring systems. The proposed method validates it’s performances across various conditions and provides output comparable to reference industrial sensors in the market as well as against state-of-the-art algorithms implemented in research to the best of our knowledge. Finally, an auto-correlation based signal periodicity computing approach was used to enable accurate separation of hidden heart-rate-related signals from noisy head-motion signals. The new method showed greater precision (i.e., smaller mean errors in heart rate measurement) and better outcomes compared with previous systems, and the precision was adequate for realistic heart rate monitoring.

References
[1] IOP Publishing is to grateful Mark A Caprio, Center for Theoretical Physics, Yale University, for permission to include the iopart-num BιβιοιΟΧπακκαγιε (version 2.0, December 21, 2006) with this documentation. Updates and new releases of iopart-num can be found on www.ctan.org (CTAN).