HEART RATE ESTIMATION FROM PPG SIGNAL USING SOFT DECISION SUBMITTED TO 2015 IEEE SIGNAL PROCESSING CUP

Students: Oz Sharlin, Aviya Maimon and Tom Srugo. Supervisors: Rami Cohen and David Malah.

Signal and Image Processing Lab (SIPL)
Department of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 3200000, Israel

ABSTRACT

In this paper, we describe an algorithm for estimating heart rate from an optically measured PPG signal when physical exercises are performed. In this case, the PPG signal is contaminated by motion artifacts caused by hand movements, making it difficult to find its fundamental frequency that corresponds to the heart rate. To overcome the noise, a soft decision approach is taken, by which several candidates for the fundamental frequency of the PPG signal are extracted and assigned grades. By appropriate grade weighting, the candidate having the maximal grade is selected. The presented algorithm is of low complexity and shown to provide good results. As such, it can be used in low-power portable devices for real time heart rate estimation.

Index Terms— PPG, motion artifacts, heart rate monitoring, signal processing cup

1. INTRODUCTION

Heart rate monitoring during physical exercise has become increasingly popular recent years. The monitoring is performed using wearable devices, which estimate heart rate in real time using photoplethysmographic (PPG) signals [1]. These PPG signals are obtained by illuminating the skin by a light-emitting diode (LED) and measuring changes in light absorption by a photodiode. As the heart pumps blood through organs, volumetric changes of organs occur, reflected in periodic variations in measured light intensity. These variations are used in turn to determine the heart rate, usually in terms of beats-per-minute (BPM).

However, physical exercises may lead to motion artifacts in the measured PPG signal, affecting its periodicity and resulting in an erroneous estimation of the heart rate. These artifacts are difficult to overcome, due to frequency overlapping between the non-contaminated PPG signal and the motion artifacts. Several techniques have been proposed for motion artifacts removal, where a literature survey can be found in [2]. In particular, a method for motion artifacts reduction based on simultaneous measurement of PPG signal and acceleration is proposed in [2], where acceleration is measured to better estimate frequency components contributed by motion artifacts.

The aim of this paper is to present an algorithm for heart rate estimation from PPG signals in presence of motion artifacts, by exploiting information obtained from measured acceleration data. The proposed algorithm is based on the use of soft decision, where grades are assigned to several BPM candidates by extracting certain features from the signal. For performance evaluation, we use the training database provided in [2]. This database contains recordings of 13 subjects performing various physical exercises, where two-channel PPG signals, three-axis accelerometer data, and one-channel ECG signal from the subject's chest are available for each subject. All signals were sampled at 125Hz. The heart rate calculated using the ECG signal serves as ground-truth for performance evaluation of the proposed algorithm. There is also a test database with 10 subjects, for which ground-truth BPM is not provided.

The paper is structured as follows. In Section 2, the algorithm is presented. Performance evaluation of the algorithm is provided in Section 3. Finally, the paper is concluded in Section 4.

2. ALGORITHM DESCRIPTION

In the following ,the proposed algorithm for heart rate estimation from PPG signals is described. As in [2], the signal is partitioned into time windows of 8 seconds with an overlapping of 6 seconds, such that an estimated BPM value is provided every 2 seconds. The algorithm has two modes of operation: *Rest mode* and *Activity mode*. Rest mode is used when a subject is at rest, whereas Activity mode is used when a physical exercise is performed. The transition between the modes is determined by the algorithm, as will be described in Section 2.2.

A block diagram describing the main parts of the algorithm is provided in Figure 1. In rest mode, the decision is made by choosing the spectral peak with the highest amplitude, out of several candidates. In activity mode, each spectral peak is assigned 5 grades, which are later weighted to provide an output. In the rest of this section, details of the algorithm components described in Figure 1 are provided.

2.1. Rest mode

In this mode, the noise is assumed to be of relatively low magnitude, affecting the PPG signal only to a small extent. Therefore, the algorithm in this mode is mainly based on recovering the periodic components of the PPG signal by analyzing its spectrum. First, the time series in each of these windows are decomposed into a sum of time series using singular spectrum analysis (SSA) [3]. This is performed as described in Section

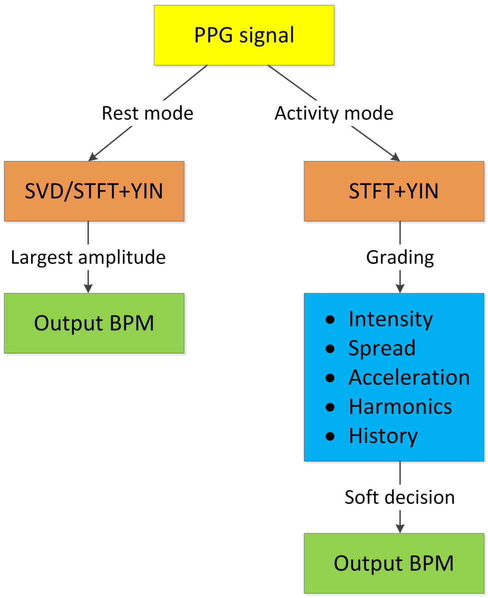


Fig. 1: Block diagram of the proposed algorithm.

Gender	Healthy?	BPM range
Male	Yes	$[45, 90]$
Male	No	$[78, 150]$
Female	Yes	$[48, 90]$
Female	No	$[78, 150]$

Table 1: Expected BPM range at rest according to gender and health status (based on [4, 5]).

III-A of [2], by constructing a matrix composed L-lagged vectors.

This matrix is decomposed using singular value decomposition (SVD), producing a basis of time series and their associated singular values. The dominant frequencies, considered as the frequency positions of the global spectral peaks of the 5 time series with the largest singular values, are calculated for each channel. Note that the corresponding BPM values are simply obtained by multiplying these frequencies by 60. In case there are no 5 feasible spectral peaks for each channel, we perform a short-time Fourier transform (STFT) on the analysis window after weighting with Hamming windows and zero-padding to $16,384$ samples to obtain additional spectral peaks. Gender and health status (with respect to cardiovascular diseases) are used to consider only physically feasible BPM values among the candidate spectral peaks. Based on [4, 5], the feasible BPM range at rest is provided in Table 1 (a tolerance of 15% is allowed).

In addition to SVD and STFT, we use YIN pitch detection algorithm [6], known to provide good results for the estimation of the fundamental frequency of speech or musical sounds. YIN is based on the autocorrelation method with several modifications, making this algorithm robust to errors. In our case, YIN is used to obtain an additional candidate spectral peak for the underlying fundamental frequency of the PPG signal. Finally, out of the extracted spectral peaks, the frequency of the peak having the largest amplitude is used for estimating the current BPM.

To ensure a physical behaviour of the estimated BPM values, we require that the estimated BPM at time window i should be within λ percent of the BPM estimated at time window $i-1$. The value of λ at rest was set experimentally at 5%. If none of the candidate spectral peaks (estimated for the two channels) are within λ percent, the previously estimated BPM value is used as the current estimated value. In this case λ is increased by 2%, to prevent drift error due to repeated use of previously estimated BPM values. λ is set again to its initial value of 5% when the current estimate falls within the current value of λ compared to the previous estimate.

2.2. Activity mode

In Activity mode, the algorithm described earlier may not be accurate due to significant motion artifacts. To determine when to move to this mode, the difference of the acceleration 2-norm of each two consecutive time windows is calculated. If a difference above 200% is detected, Activity mode is selected. The algorithm keeps track of the acceleration norm, for detecting transitions to a less/more extensive physical exercise. This is done by defining a ternary *trend* parameter ρ, whose values can be either $-1, 0$ or 1. If the average norm difference in the last 8 time windows is above (below) a threshold ρ_T $(-\rho_T)$, ρ is set to 1 (-1). Otherwise, ρ is set to 0. ρ_T was determined experimentally as 0.0012 for the training set and as 0.005 for the test set, then Activity mode is used.

Each time window is weighted using Hamming window and its spectrum is produced by applying the STFT. 5 global spectral peaks with the largest amplitude (for each channel) are selected from the spectrum amplitude along with an additional estimate of the fundamental frequency provided by the YIN algorithm. These 12 frequency values (6 for each channel) serve as candidates for the estimated BPM. Note that SVD is not used in this mode, since experimental results showed no significant improvement over the use of frequency values obtained using STFT and YIN when motion artifacts are presented. This implies lower complexity of the algorithm when physical exercises are performed, making it attractive for real-time implementation.

Denote by $f_{i,j}$ the j^{th} $(j = 1, 2, ..., 6)$ frequency position (in Hz) of the candidate spectral peaks selected for channel i $(i = 1, 2)$. In Activity mode, we consider only $f_{i,j}$ values corresponding to BPM values smaller than $217 - 0.85 \cdot$(Subject age) [7] (with a tolerance of 10%). To choose the most likely candidate, we assign *grades* to each spectral peak frequency position. In this grading process we take into account features associated with each $f_{i,j}$, together with information regarding the noise, which is provided by the 3-axis accelerometer data. There are 5 grades for each $f_{i,j}$, denoted $g_{i,j}^{(t)}$, where t denotes the grade index $(t = 1, 2, ..., 5)$. The grades are calculated as follows.

1. **Intensity**. The intensity of $f_{i,j}$ is defined as its amplitude (in the STFT domain), denoted $A_{i,j}$. The *relative intensity* of $f_{i,j}$ is obtained by dividing $A_{i,j}$ by the maximum amplitude obtained for channel i. The corresponding grade of $f_{i,j}$, denoted $g_{i,j}^{(1)}$, is:

$$g_{i,j}^{(1)} = \frac{A_{i,j}}{\max_j A_{i,j}}, \qquad i = 1, 2. \tag{1}$$

Note that $A_{i,j}$ is normalized using the maximum amplitude obtained for channel i only, due to possible scale variations between channels. According to Equation (1), $f_{i,j}$ having a higher amplitude is assigned a higher grade.

2. **Spread**. The spread of $f_{i,j}$, denoted $A_{i,j}$, is defined as the energy contained in the the spectrum band centered at $f_{i,j}$. The bandwidth is defined as the distance of $f_{i,j}$ to the nearest local minima. For better estimate of the spread, a second order polynomial is fitted to the spectrum amplitude at $f_{i,j}$ assuming bandwidth as above. To calculate the grade, we subtract from 1 the normalized spread of $f_{i,j}$:

$$g_{i,j}^{(2)} = 1 - \frac{S_{i,j}}{\max_j S_{i,j}}, \qquad i = 1, 2. \tag{2}$$

This way, a lower spread results in a higher grade.

3. **Acceleration**. This grade is calculated as the normalized minimal distance of $f_{i,j}$ from the 2 dominant frequency components of the acceleration signal (for each axis). The underlying assumption is that these components are related to motion artifacts, such that $f_{i,j}$ is likely to be noise if it is close to the acceleration dominant frequencies. Denoting the acceleration frequency components by $a_{k,l}$ ($k = 1, 2, 3$ refers to the axis index and $l = 1, 2$ refers to the frequency index), the grade of $f_{i,j}$ is calculated as:

$$g_{i,j}^{(3)} = \frac{\min_{k,l} |f_{i,j} - a_{k,l}|}{\max_{i,j} \left(\min_{k,l} |f_{i,j} - a_{k,l}|\right)}. \tag{3}$$

4. **Harmonics**. As observed in [8], the motion artifacts have higher harmonic content compared to the PPG signal. Therefore, $f_{i,j}$ will be given a lower grade if there are other frequency positions which are integer multiples of $f_{i,j}$ (a tolerance of 25% is allowed). If there are no such integer multiples, $g_{i,j}^{(4)}$ is set to 1. Otherwise, denote by $h_{k,l}$ the detected harmonics of $f_{i,j}$. The grade in this case is calculated as the sum of distances to the harmonics, normalized by the number of harmonics:

$$g_{i,j}^{(4)} = \frac{\sum_l |f_{i,j} - h_{k,l}|}{\# h_{k,l}}. \tag{4}$$

This way, as the harmonics $h_{k,l}$ are close to $f_{i,j}$ and as their number grows, $g_{i,j}^{(4)}$ will be lower, since $f_{i,j}$ will be attributed as noise.

5. **History**. The aim of this grade is to ensure physical behaviour of the estimated BPM. That is, if $f_{i,j}$ is not within λ percent of the previously estimated signal, its grade would be $g_{i,j}^{(5)} = 0$. Otherwise, its grade is calculated as the normalized distance to the previously estimated heart rate frequency. Since heart rate variation is higher when a physical exercise is performed, we set $\lambda = 12\%$ in Activity mode. In addition, we use the trend parameter ρ for improving the prediction. For example, if some $f_{i,j}$ is larger than the previously estimated frequency by less than λ percent and $\rho = 1$, then its history grade is doubled. If $\rho = -1$, then the grade is reduced by a factor of 2.

Finally, we choose the $f_{i,j}$ with the maximal weighted grade, and calculate the estimated BPM value by multiplying it by 60. Experimentally, the grade weights were set to $0.6, 0.8, 0.8, 0.4, 1$ for grades $1, 2, 3, 4, 5$, respectively.

3. EXPERIMENTAL RESULTS

The performance of the proposed algorithm was evaluated using the training database, by comparing the estimated BPM to the ground-truth provided by the ECG signal. Denote the estimated BPM value at time window i by $\text{BPM}_{est}(i)$, the ground-truth BPM by $\text{BPM}_{true}(i)$, and the number of time windows by N. The measurement index we pursue here is the average absolute error, defined as [2]:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} |\text{BPM}_{est}(i) - \text{BPM}_{true}(i)|. \tag{5}$$

The results for the training set are provided in Table 2, using the error definition of Equation (5). For comparison, the results of [2] are provided as well, where the best case for each subject is marked in bold. The results show the overall good performance of the proposed algorithm, with an average error of 3.04. Taking into account the first 12 subjects only (as in [2]), the average error of the proposed algorithm is 2.85, compared to 2.34 in [2]. Note that the proposed algorithm is significantly less complex than the algorithm presented in [2], in which SSA is performed for each time window and sparse signal reconstruction is required.

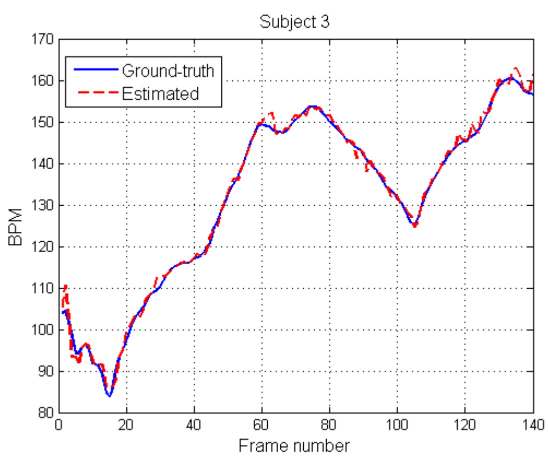
The algorithm operation is demonstrated in Figure 2 for Subject 3 and Subject 13 from the training set. Good BPM estimate is obtained for Subject 3 (Figure 2a), since his ground-truth BPM curve is relatively smooth, without abrupt changes. However, the ground-truth BPM curve of Subject 13 (Figure 2b) changes fast, with abrupt changes even when the same physical exercise is performed (likely indicating Arrhythmia). This makes the estimation process difficult, such that BPM changes are not well captured. It is expected that the performance of the algorithm for the test data would be more accurate in tracking BPM changes, since gender/age/health status are provided for the data set subjects. Examples for BPM estimation results for the test set are provided in Figure 3.

4. CONCLUSION

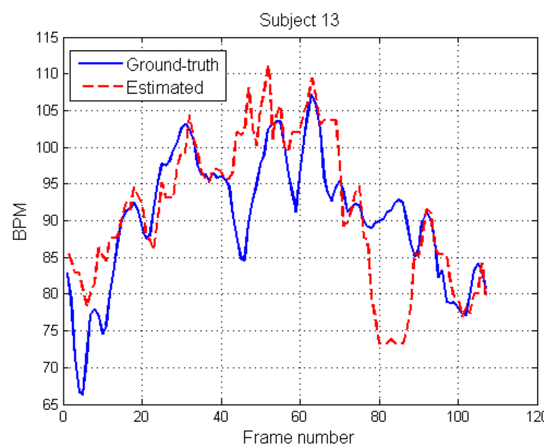
In this paper, we proposed an algorithm for estimating heart rate from a dual-channel PPG signal, when physical exercises are performed. Based on the selected mode (Rest/Activity) of the algorithm, it provides several candidates for the heart rate at each time interval. The most likely candidate is selected based on features extracted from the signal and on previously estimated BPM values, by assigning a weighted grade to each candidate. Information regarding age, gender and health status is incorporated in the decision process for improved accuracy. The proposed algorithm is of low-complexity and can be efficiently implemented in hardware. The algorithm is modular as well, where grades can be omitted/added according to the allowed complexity when real-time constraints are considered.

Method/Subject	1	2	3	4	5	6	7	8	9	10	11	12	13
Proposed algorithm	**2.22**	**1.26**	**0.91**	**1.69**	2.64	3.00	**1.10**	4.31	6.29	6.02	2.27	**2.50**	5.43
TROIKA [2]	2.29	2.19	2.00	2.15	**2.01**	**2.76**	1.67	**1.93**	**1.86**	**4.70**	**1.72**	2.84	–

Table 2: Average absolute error on the training database.

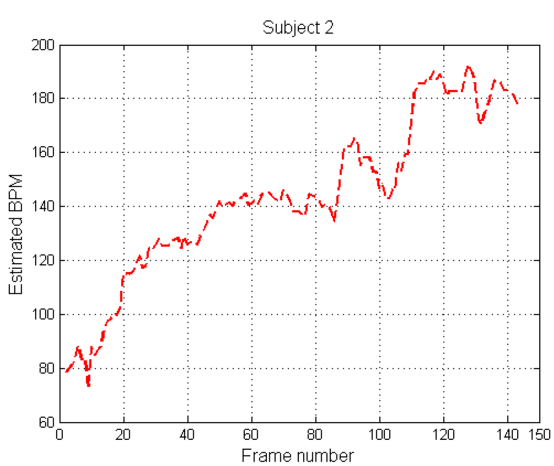


(a) Subject 3.

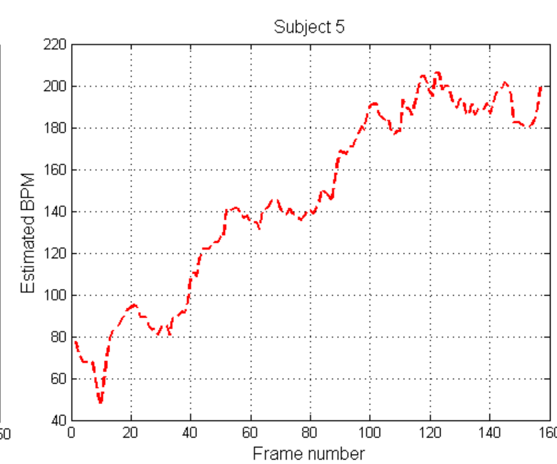


(b) Subject 13.

Fig. 2: Examples of training set BPM estimation results.



(a) Subject 2.



(b) Subject 5.

Fig. 3: Examples of test set BPM estimation results.

5. REFERENCES

[1] John Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiological Measurement*, vol. 28, no. 3, pp. R1–R39, February 2007.

[2] Zhilin Zhang, Zhouyue Pi, and Benyuan Liu, “TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 522–531, February 2015.

[3] Nina Golyandina, Vladimir Nekrutkin, and Anatoly A. Zhigljavsky, *Analysis of Time Series Structure: SSA and related techniques*, CRC Press, 2001.

[4] Ariel Diaz, Martial G. Bourassa, Marie-Claude Guertin, and Jean-Claude Tardif, “Long-term prognostic value of resting heart rate in patients with suspected or proven coronary artery disease,” *European Heart Journal*, vol. 26, no. 10, pp. 967–974, March 2005.

[5] Hirofumi Tanaka, Kevin D Monahan, and Douglas R Seals, “Age-predicted maximal heart rate revisited,” *Journal of the American College of Cardiology*, vol. 37, no. 1, pp. 153–156, 2001.

[6] Alain de Cheveigne and Hideki Kawahara, “YIN, a fundamental frequency estimator for speech and music,” *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1917–1930, April 2002.

[7] Robert A. Robergs and Roberto Landwehr, “The suprising history of the HRmax=220-age equation,” *Journal of Exercise Physiology*, vol. 5, no. 2, pp. 1–10, May 2002.

[8] S. M. Lopez-Silva et al., “Heuristic algorithm for photoplethysmographic heart rate tracking during maximal exercise test,” *Journal of Medical and Biological Engineering*, vol. 32, no. 3, pp. 181–188, January 2012.