PPG-BASED HEART RATE ESTIMATION USING WIENER FILTER, PHASE VOCODER AND VITERBI DECODING

Andriy Temko

Irish Center for Fetal and Neonatal Translational Research, University College Cork, Ireland

ABSTRACT

Accurate HR estimation from the photoplethysmographic (PPG) signal during intensive physical exercises is tackled in this paper. Wiener filters are designed to attenuate the influence of motion artifacts. The phase vocoder is used to improve the initial DFT-based frequency estimation. Additionally, Viterbi decoding is used as a novel post-processing step to find the path through time-frequency state-space plane. The system performance is assessed on a publically available dataset of 23 PPG recordings. The resulting algorithm is designed for scenarios that do not require online HR monitoring (swimming, offline fitness statistics). The resultant system with an error rate of 1.37 beats per minute outperforms all other systems reported to-date in literature and in contrast to existing alternatives requires no parameter to tune at the post-processing stage and operates at a much lower computational cost. The Matlab implementation is provided online.

***Index Terms*—**Photoplethysmography, motion artifacts, spectrum estimation, Viterbi decoding

1. INTRODUCTION

Wearable devices such as wrist-bands, smart-watches, are equipped with a number of sensors and offer many useful fitness tracking features. Photoplethysmography (PPG) based heart rate (HR) monitoring has become a popular alternative to traditional Electrocardiography as it allows for HR monitoring at the peripheral positions such as earlobes, fingertips or wrists which is seen much more convenient [1].

PPG sensors which are embedded in these wearable devices emit light to the skin and measure the changes of intensity of the light which is reflected or transmitted through the skin. The periodicity of these measurements in most cases corresponds to the cardiac rhythm, and thus, HR can be estimated from the PPG signal [1, 2].

During physical exercise the PPG signal is corrupted with motion artifacts (MAs). MAs significantly affect the accuracy of HR estimation in free living conditions thus preventing the straight-forward usage of PPG. A number of methods have been proposed to remove or attenuate MAs in PPG signals using the simultaneously recorded accelerometer signals. These methods include adaptive filtering [3, 4, 5, 6], independent component analysis [7], decomposition models [8, 9, 10, 11], spectral subtraction [12, 13, 14], and Kalman filtering [15].

Along with de-noising routines the accurate methods utilize sophisticated post-processing of HR estimates [2, 6, 8, 10, 11, 14]. The post-processing steps often include spectral peak detection, peak selection, temporal peak tracking, smoothing, etc. The post-processing is usually composed of several heuristic *if-then* rules and associated thresholds. These thresholds are tuned and tested on the same data. However, when reporting the results, the effect of post-processing is often overlooked.

This work enhances the previously developed HR estimation system [16] with a threshold-free post-processing step. Specifically, a time-frequency spectrogram is seen as a state-space matrix of emission probabilities and the Viterbi decoding algorithm is used to find the most probable path through the PPG recording.

2. DATABASE AND METRICS

The dataset which is publically available (http://zhilinzhang.com/spcup2015/data.html) consists of 23 5-min recordings of subjects performing various physical exercises ranging from walking or running on a treadmill (recordings 1-12) to jump-ups or boxing (recordings 13-23). Each recording consists of two PPG and three accelerometer signals. The ECG signal which was recorded simultaneously from the chest was used to provide the ground truth HR in BPM as described in detail in [2, 8]. ECG-based HRs were calculated for every 8s window with a 2s shift. The same window length and shift are suggested for HR estimation from PPG to have the same number of HR estimates and true HRs. All signals were sampled at 125 Hz.

The conventional metrics to measure the performance of developed HR estimators are based on the Absolute Error (AE) of each estimate:

This research was supported by a Science Foundation Ireland Centers Award (12/RC/2272) and Wellcome Trust Seed Award in Science (200704/Z/16).

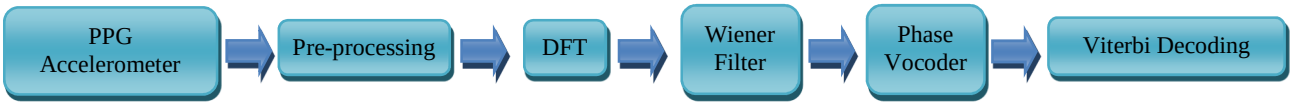

Fig.1. The flowchart of the developed HR estimation system (WFPV+VD).

$$AE_i = \left| BPM_{est}(i) - BPM_{true}(i) \right| \quad (1)$$

where $BPM_{est}(i)$ and $BPM_{true}(i)$ denote the estimated and the true HR value in the i-th time window in BPM, respectively. To summarize the performance for a whole recording, Average Absolute Error (*avAE*) and Standard Deviation of the Absolute Error (*stdAE*) are reported:

$$avAE = \frac{1}{N}\sum_{i=1}^{N} AE_i \quad (2)$$

$$stdAE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(AE_i - avAE\right)^2} \quad (3)$$

where N is the total number of estimates (number of windows).

These metrics are computed for each of the 23 recordings and the average performance across subjects is reported.

3. HR ESTIMATOR DESIGN

The flowchart of the developed system (WFPV+VD) is shown in Fig. 1. The PPG and accelerometer signals are segmented to 8s windows with 2s shift, filtered with a 4th order Butterworth band-pass filter (0.4-4Hz) and down-sampled from 125 to 25 Hz as shown in Fig. 2(a). The PPG signals are then normalized to zero mean and unit variance and averaged. The signals are then subjected to the DFT with the number of bins set to 1024. The content that corresponds to the HR between 60 BPM and 180 BPM is kept as shown in Fig. 2(b).

Wiener filtering [17] is applied to attenuate the effect of MA in the PPG signal. The frequency-domain Wiener filter is given as:

$$W(f) = \frac{P_{XX}(f)}{P_{XX}(f) + P_{NN}(f)} \quad (4)$$

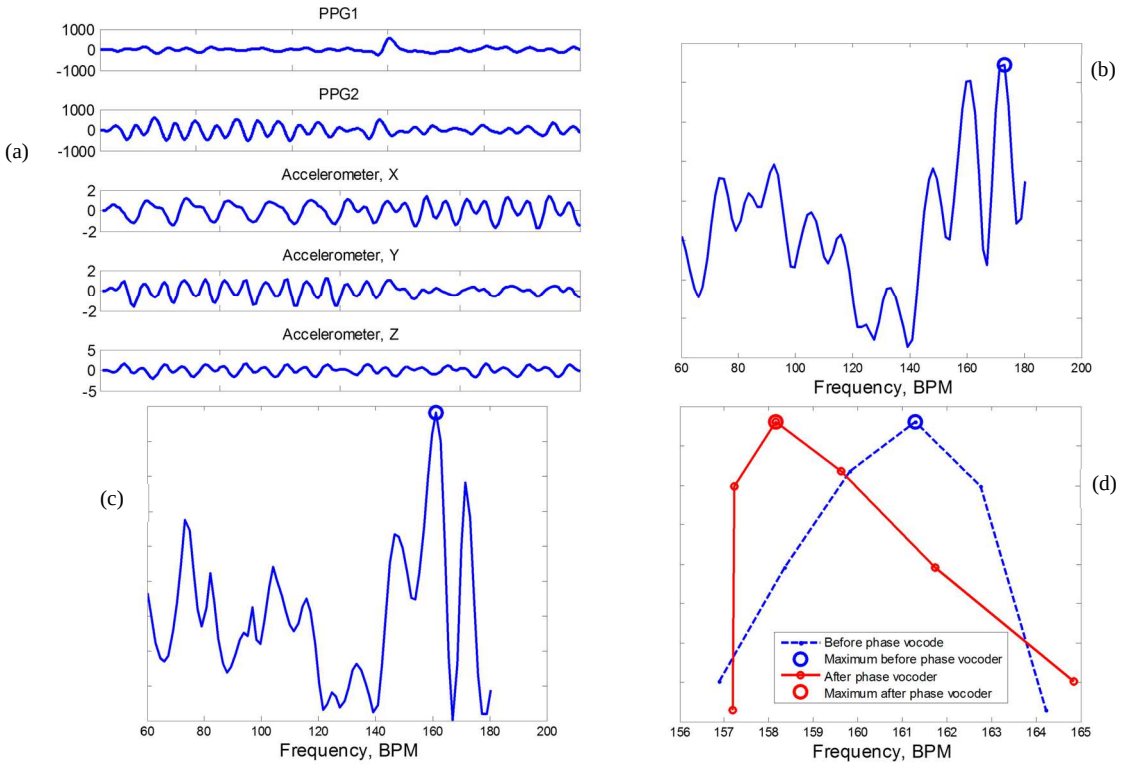

Fig. 2. Signal transformation in the developed HR estimation system. Plot (a) shows two PPG and three accelerometer signals after filtering; (b) shows the spectral envelope and its maximum after the DFT is applied to the PPG, (c) shows the processed spectral envelope and its maximum after MAs were attenuated with Wiener filtering; (d) shows the maximum of the spectral envelop before and after the phase vocoder.

The noise spectrum, P_{NN}, is estimated from the accelerometer signals. The clean PPG spectrum, P_{XX}, can be estimated as $P_{YY}(f) - P_{NN}(f)$ or recursively from previous filter outputs. Depending on how the power spectrum of the clean PPG signal is estimated, two Wiener filters are implemented:

$$w_1(t,k) = \frac{P_{YY}(t,k) - P_{NN}(t,k)}{\frac{1}{C_1}\sum_{i=t-C_1+1}^{t} P_{YY}(i,k)} \quad (5)$$

$$w_2(t,k) = \frac{\frac{1}{C_2}\sum_{i=t-C_2}^{t} P_{YY}(i,k)\, w_2(i,k)}{\frac{1}{C_2}\sum_{i=t-C_2}^{t} P_{YY}(i,k)\, w_2(i,k) + P_{NN}(t,k)} \quad (6)$$

where $w(t,k)$ is the weight of the k-th frequency bin at time, t. The power spectrum of the observed signal is averaged over the past C spectral envelopes (C_1=1, C_2=3). If C=0, then the Wiener filter in Eq. 5 performs a simple version of spectral subtraction.

The outputs of both Wiener filters are normalized by their standard deviation and averaged to represent the final spectral envelope of the cleaned PPG signal. The dominant frequency (the frequency with the highest magnitude) is converted to the HR estimate in BPM as shown in Fig. 2(c).

The minimum frequency that can be estimated (the Rayleigh frequency) is limited by the size of the window of the analyzed data (8s) and equals to 1/8*60 = 7.5 BMP. Apart from zero-padding which is used to interpolate the spectral envelope to other frequencies by decreasing the frequency spacing between neighboring DFT bins, the frequency estimate is improved by means of the phase vocoder technique [18–20]. Phase changes between two consecutive frames encode the deviation of the true frequency from the bin frequency. Thus the phases from the chosen peak in the magnitude spectrum from the current and previous frames can be used to refine the initial frequency estimation:

$$\phi_{new} = (\theta_2 - \theta_1 + 2\pi n)/\big(2\pi(t_2 - t_1)\big) \quad (7)$$

where θ_2, θ_1 are the two phases from the current and previous frames, respectively; t_2, t_1 are the time stamps of the two frames, here t_2-t_1 is a windows shift and is equal to 2s, n is an integer. The values of ϕ_{new} are computed for several n using Eq. 7, and the value that is closest to the initial frequency estimation is chosen. In this manner the dominant frequency peak is adjusted accordingly as shown in Fig. 2(d).

The novel step proposed in this work performs post-processing in a probabilistic framework using Viterbi decoding [21]. The spectrogram of a cleaned PPG recording (after Weiner filtering) is considered as a N-by-T state-space map of emission probabilities (likelihoods), B, for N states (discrete values of HR as determined by the size of DFT) and

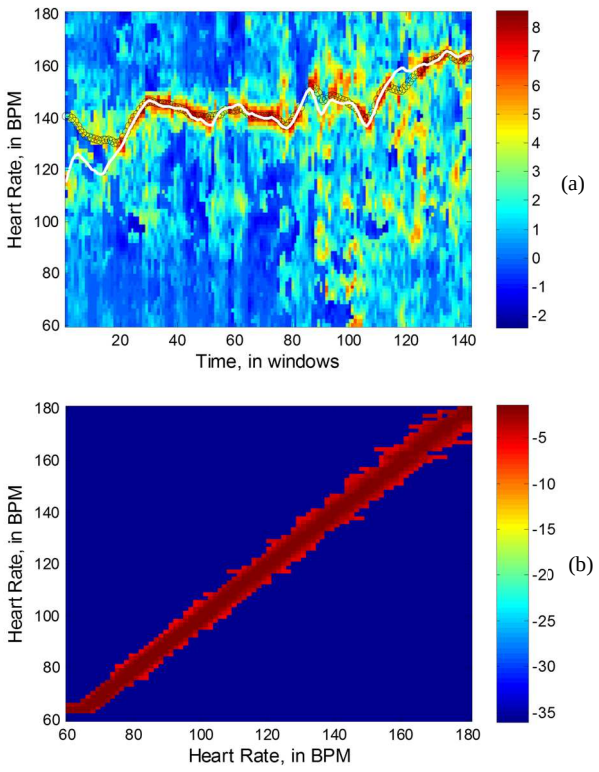

Fig. 3. (a) Time-frequency state-space plane of DFT magnitudes considered as log emission probabilities, with superimposed ground truth HRs in white. The output of Viterbi decoding is shown in black. (b) a matrix of estimated log transition probabilities. Best viewed in colour.

T observations (time windows), where B_{jt} is a magnitude value of the j^{th} DFT bin for the t^{th} time window. An example of emission probability matrix is shown in Fig. 3(a) for rec. 21. The N-by-N matrix of transition probabilities, A, where A_{ij} represents the probability of changing from the i^{th} HR to the j^{th} HR, is estimated from the ground truth by counting the transitions using the leave-one-recording-out procedure. In this manner, the ground truth of the testing recording is never used but the ground truths of all other recordings are used to estimate the transition probability matrix. An example of a log transition probability matrix is visualized in Fig. 3(b). It can be seen that the variance of the transition probabilities increases with the increase of HR.

The Viterbi algorithm is then applied to recursively estimate the most likely path (the path with the highest cumulative log probability) through the time dimension, t, using emission and transition probability matrices, B, A:

$$\Psi_t(j) = \underset{1\le i\le N}{\operatorname{argmax}}\big[\delta_{t-1}(i)A_{ij}\big], \quad (8)$$

where

$$\delta_t(j) = \max_{1\le i\le N}\big[\delta_{t-1}(i)A_{ij}\big]B_{jt}, \quad 2 \le t \le T, \text{and } 1 \le j \le N,$$

and $\delta_1(i) = \pi_i B_{i1}, \quad 1 \le i \le N$

The prior probabilities, π_i, are approximated as $diag(A)$. After the recursion is computed the state sequence is backtracked as:

$$i_T = \underset{1 \le i \le N}{\operatorname{argmax}}[\delta_T(i)], \qquad (9)$$
$$i_t = \Psi_{t+1}(i_{t+1}), t = T-1, T-2, \ldots, 1$$

The state sequence is converted to HR estimates which are then smoothed with a central moving average filter. The Viterbi decoding effectively performs post-processing in a threshold-free probabilistic manner.

4. RESULTS AND DISCUSSION

Table I details the performance of the system. On the database of 23 recordings, the system results in an *avAE* of 1.31 BPM with *sdAE* of 1.77 BPM.

Table I also shows the performance of the system if a certain block from Fig. 1 is removed. This gives an indication of the contribution of each system constituent towards the final performance. It can be seen that without Wiener filtering (W/o WF in Table I), the performance degrades from 1.31 to 5.71 BPM. This result represents a HR estimation system that does not address MAs and also shows the effect of MAs in the database. Using only WF1 results in an *avAE* of 1.43 BPM while using only WF2 results in an *avAE* of 1.46 BPM, to compare with the *avAE* of 1.31 BPM using both filters combined. A refinement introduced by the phase vocoder reduces the error from 1.47 to 1.31 BPM (W/o PV in Table I). Finally, the system performance without Viterbi post-processing (W/o VD) results in an *avAE* of 5.86 BPM.

It can be seen that post-processing of HR estimations has a significant effect on the performance, similar in size to the contribution of signal de-noising. The post-processing steps in approaches which were previously evaluated on the same dataset usually rely on a number of heuristic rules and thresholds [2, 8, 6, 10, 11, 14]. These rules and threshold values are tuned and tested on the same data and the best possible results are usually reported. In this manner, the number of rules (degrees of freedom) is directly linked to an improved performance but it comes at the cost of an increased risk of poor generalization on the unseen data. It is worth emphasizing that in this work the proposed post-processing requires no thresholds and is performed in the well-established probabilistic framework.

Table II provides a comparison with other HR estimation algorithms. Many developed systems report results only on an 'easier' part of the dataset (the first 12 PPG recordings) [6, 10, 11]. These PPGs are captured during running on a treadmill. The recordings are less corrupted with irregular MAs. Only a few techniques are evaluated on a more difficult part of the dataset (the last 10 recordings), in addition to the first 12. This is replicated here for comparative purposes (that is the WFPV+VD is evaluated on 22 recordings excluding rec. 13). These systems include: a three-stage method which is based on signal decomposition, sparsity-based high-resolution spectrum estimation, and spectral peak tracking and verification (TROIKA, [2]); a method which jointly estimates the spectrum of PPG and accelerometer signals using a common sparsity constraint on the spectral coefficients (JOSS, [8]); a time-varying spectral filtering algorithm for reconstruction of motion artifact (SpaMa, [13]), a HR estimation algorithm based on asymmetric least squares spectrum subtraction and Bayesian decision theory (Spectrap, [14]). As it can be seen from Table II, the WFPV+VD system outperforms every method on the database. It is worth mentioning that the first three methods perform online processing, whereas Spectrap is the only method that performs offline processing. Viterbi decoding in this work requires the spectrogram of the whole recording to be available beforehand.

TABLE I.
PERFORMANCE OF HR ESTIMATION SYSTEM WFPV+VD ON 23 PPG RECORDINGS

		WFPV+VD	W/o WF	W/o PV	W/o VD
All	***avAE***	**1.31**	5.71	1.47	5.86
	sdAE	**1.77**	7.16	1.81	9.10

TABLE II.
COMPARISON OF HR ESTIMATION SYSTEMS ON 22 PPG RECORDINGS

	TROIKA	JOSS	SpaMa	Spectrap	WFPV+VD
avAE	2.73	2.08	2.01	1.79	**1.24**

The proposed algorithm takes under 10s to process the whole PPG dataset of 23 recordings (Matlab R2013b @ Intel Core E7200 2.5GHz). This time compares favorably with other techniques published. It is reported that to process the first 12 PPG recordings TROIKA [2] and IMAT [6] takes several hours, JOSS [8] takes 300s, EEMD [10] takes 200s.

The presented system with its superior accuracy and a comparatively low computational cost can be used for HR monitoring for swimmers or calculation of HR summary statistics. Both tasks require quick processing but do not require on-the-fly HR estimation.

For the purpose of reproducibility the Matlab implementation along with the main results is available online (https://github.com/andtem2000/PPG).

5. CONCLUSIONS

This work proposed an alternative threshold-free step to post-process HR estimates which were derived from PPG. It was shown that HR estimate post-processing and motion artefact rejection had a commensurable effect on the performance. Viterbi decoding was applied to spectrogram to find the path through time-frequency state-space plane of HR estimates. Evaluated on a publicly available dataset the system resulted in an error of 1.37 BPM which outperformed alternatives reported to-date. The novel post-processing requires no parameter to tune and resultant system operates at a much lower computational cost.

6. REFERENCES

[1] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Phys Meas*, v. 28, pp. 1-39, 2007.

[2] Z. Zhang, Z. Pi, B. Liu, "TROIKA: A General Framework for Heart Rate Monitoring Using Wrist-Type Photoplethysmographic Signals During Intensive Physical Exercise," *IEEE Trans Biomed Eng*, v. 62, pp. 522-531, 2015.

[3] R. Yousefi, M. Nourani, S. Ostadabbas, and I. Panahi, "A motion-tolerant adaptive algorithm for wearable photoplethysmographic biosensors," *IEEE J Biomed Health*, v. 18, pp. 670–681, 2014.

[4] M. Ram, K. V. Madhav, E. H. Krishna, N. R. Komalla, and K. A. Reddy, "A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter," *IEEE Trans Instrum Meas*, v. 61, pp. 1445–1457, 2012.

[5] H. Pan, D. Temel, G. AlRegib, "HeartBEAT: Heart beat estimation through adaptive tracking," in Proc. IEEE EMBC, 2016.

[6] M. Mashhadi, E. Asadi, M. Eskandari, S. Kiani, F. Marvasti, "Heart Rate Tracking using Wrist-Type Photoplethysmographic (PPG) Signals during Physical Exercise with Simultaneous Accelerometry", *IEEE Signal Processing Letters*, v. 23, 2016.

[7] B. Kim, S. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *IEEE Trans Biomed Eng*, v. 53, pp. 566–568, 2006.

[8] Z. Zhang, "Photoplethysmography-Based Heart Rate Monitoring in Physical Activities via Joint Sparse Spectrum Reconstruction," *IEEE Trans Biomed Eng*, v.62, pp. 1902-1910, 2015.

[9] X. Sun, P. Yang, Y. Li, Z. Gao, and Y.-T. Zhang, "Robust heart beat detection from photoplethysmography interlaced with motion artifacts based on empirical mode decomposition," in Proc. IEEE BHI, pp. 775–778, 2012.

[10] E. Khan, F. Al Hossain, S. Uddin, S. Alam, M. Hasan, "A Robust Heart Rate Monitoring Scheme Using Photoplethysmographic Signals Corrupted by Intense Motion Artifacts," *IEEE Trans Biomed Eng*. v.63, pp. 550-562, 2016.

[11] J. Xiong, L. Cai, D. Jiang, H. Song, X. He, "Spectral Matrix Decomposition-Based Motion Artifacts Removal in Multi-Channel PPG Sensor Signals," *IEEE Access*, 2016.

[12] H. Fukushima, H. Kawanaka, M. Bhuiyan, and K. Oguri, "Estimating heart rate using wrist-type photoplethysmography and acceleration sensor while running," in Proc. IEEE EMBC, pp. 2901–2904, 2012.

[13] S. Salehizadeh, D. Dao, J. Bolkhovsky, C. Cho, Y. Mendelson, K. Chon, "A Novel Time-Varying Spectral Filtering Algorithm for Reconstruction of Motion Artifact Corrupted Heart Rate Signals During Intense Physical Activities Using a Wearable Photoplethysmogram Sensor," *Sensors*, 16(1), 2016.

[14] B. Sun, Z. Zhang, "Photoplethysmography-based heart rate monitoring using asymmetric least squares spectrum subtraction and bayesian decision theory," *IEEE Sensors J*., v. 15, pp.7161-7168, 2015.

[15] B. Lee, J. Han, H. Baek, J. Shin, K. Park, and W. Yi, "Improved elimination of motion artifacts from a photoplethysmographic signal using a kalman smoother with simultaneous accelerometry," *Physiol Meas*, vol. 31, no. 12, pp. 1585–1603, 2010.

[16] A. Temko, "Estimation of Heart Rate from Photoplethysmography during Physical Exercise using Wiener Filtering and the Phase Vocoder," in Proc. IEEE EMBC, Aug. 2015.

[17] R. Brown, P. Hwang. Introduction to Random Signals and Applied Kalman Filtering (3 ed.). New York: John Wiley & Sons, 1996.

[18] J. Flanagan, R. Golden, "Phase Vocoder," Bell System Technical Journal, 1493-1509, 1966.

[19] T. Dutoit, F. Marquès. Applied Signal Processing – A MATLAB-Based Proof of Concept. Springer, 2009.

[20] A. Temko, W. Marnane, G. Boylan, G. Lightbody, "Clinical Implementation of a Neonatal Seizure Detection Algorithm," *Decis Support Syst*, v.70, pp. 86–96, 2015.

[21] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," Proc. of the IEEE, v. 77, pp. 257–286, 1989.