STORM: A General Framework for Heart Rate Monitoring Using Wrist-Type Photoplethysmographic (PPG) Signals During Intensive Physical Exercise

Howoong Jun, Sangil Lee and Dongsuk Sung

Department of Electric and Computer Engineering
Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul, Korea
{prestoxic & sangil07 & eastein}@snu.ac.kr

Kyogu Lee and Jaehun Kim

Music and Audio Research Group,
Graduate School of Convergence Science and Technology,
Seoul National University, Korea
{kglee & eldrin}@snu.ac.kr

***Abstract* – Heart rate monitoring using wrist-type Photo-PlethysmoGraphic (PPG) signals during subjects' intensive physical exercise has been studied a lot. However, precedent studies were not accurate for the intensive physical exercise signals. In this work, a general framework, termed STORM, is based on TROIKA (Zhinlin Zhang, 2014), which consists of Singular spectrum analysis for denoising, sparse signal reconsTruction for high-resOlution spectrum estimation, and spectRal peak tracking with seven algorithMs for voting and verification. The STORM framework uses several fundamental frequency pitch tracking algorithms in spectral peak tracking part, so it has high accuracy and estimated hear rate is robust to any kinds of motion artifacts.**

***Index Terms* – Photoplethysmograph (PPG), Singular Spectrum Analysis, Sparse Signal Reconstruction, Fundamental Frequency Pitch Tracking Algorithm, Autocorrelation, Harmonic Product Spectrum, Heart Rate Monitoring, YIN, Kalman Filtering.**

I. INTRODUCTION

The real-time heart rate (HR) monitoring is important not only for the people doing cardiovascular workout, but also for most of the people who want to manage their health. Thus, many signal-processing experts have been trying to generate the accurate HR signal using photoplethysmographic (PPG) signals. However, because of subjects' movement and unexpected errors of wearable devices, it is difficult to predict accurate HR. Fortunately, subjects' movement data, which is termed motion artifacts (MA), can be measured and removed. Therefore it is available to achieve the robust estimated HR from the MA removed PPG signals.

Our framework, which is termed STORM, is based on Zhinlin Zhang's TROIKA [1] and it is composed of three parts; Singular Spectrum Analysis (SSA), Sparse Signal Reconstruction (SSR) and Spectral Peak Track (SPT). SSA is one of the signal decomposition method, separates the MA data on the PPG signal and track proper stream from the harmonic and previous estimated data. SSR enhances the separated signal from SSA to clarify the PPG signal. Finally, SPT makes the final decision and connects the signal. The most distinguishing portion of this framework is SPT.

The rest of the paper is organized as follows. Section II presents the STORM framework. Section III introduces the result and experimental settings. Finally, Section IV concludes the paper.

II. THE STORM FRAMEWORK

The STORM framework consists of three key parts: Singular Spectrum Analysis (SSA), Sparse Signal Reconstruction (SSR) and Spectral Peak Track (SPT).

A. Singular Spectrum Analysis (SSA)

SSA is mostly based on the SSA in TROIKA [1]. It includes five steps: Embedding, Singular Value Decomposition (SVD), Grouping & Reconstruction, Excluding MA from reconstructed data, Tracing a right direction through harmonic data and past value. Embedding step to Excluding MA from reconstructed data steps are same as TROIKA and we add tracing steps. There are two ways of tracing. First, it uses the harmonics to determine whether the algorithm is tracing a proper stream or harmonic stream. If algorithm is tracing harmonic stream it moves to the proper stream. Second, it uses past estimated HR to determine whether the present stream is proper stream. If it keeps detecting the estimated HR outside of the fitting range, it jumps to the other stream where it keeps detecting.

B. Sparse Signal Reconstruction (SSR)[4]

SSR is used in variety of field because of its remarkable signal reconstruction availability. As TROIKA [1] this framework also use FOCUSS algorithm [4] in SSR section. FOCUSS algorithm is based on sparse solution. The basic SSR expression is expressed as follows,

$$y = \Phi x + v$$

where Φ is a known basis matrix of the size $M \times N$. In FOCUSS algorithm it constructing the matrix Φ as (m,n)-th component which is

$$\Phi_{m,n} = e^{j\frac{2\pi}{N}mn} \quad (m = 0, \cdots, M-1, n = 0, \cdots, N-1)$$

The vector y is an observed signal which in here we used PPG signal after SSA and temporal difference operation. v is an

TABLE 1 Weight Parameters

Spectral Peak Selection	HPS	Simple Autocorrelation	Revised YIN	NSSA	YBLS	Global Maximum
4.5333	1.0391	1.1124	0.6961	0.9330	0.7172	0.9689

unknown noise vector such as MA in this situation. From y and Φ it tries to figure out which the vector x is.

C. Blurring

In Blurring it blurs the range around the MA location by Gaussian Blurring. One of the result variables in SSA is locations of MA and it is used to set the range of blurring. The range of blurring is ($MA\, location \pm 10$) for each MA location.[1] Then PPG signals, which are reconstructed by SSA, in that blurring range, become more smooth and weaker with Gaussian blurring of standard deviation 3.

D. Spectral Peak Track (SPT)

Spectral Peak Track (SPT) is an essential part in STORM. SPT receives the raw data that is filtered out from the previous steps, SSA and SSR, and makes the final decision. SPT include two steps, voting and verification. In voting step, SPT uses seven algorithms to make result from the Spectral Peak Selection, Harmonic Product Spectrum (HPS), Simple Autocorrelation, YIN, Nssa (raw data from SSA without verification), Y_{BLS} (SSA value after verification), Global Maximum and SHARP. After all algorithms are done, SPT votes with the results from seven algorithms and find out most appropriate answer. In verification step, SPT verifies whether the present result is as similar as the past result or not.

1) Fundamental Frequency Pitch Tracking Algorithm:

Spectral Peak Selection: In the spectral peak selection function, it finds the peaks from the SSR data. After each peak is selected, it products the value of each peak position and the first harmonic position, which is double of the original peak position. Let the product multiplying peak Product of Harmonics and its maximum position Product Maximum (PM). If PM is between the minimum peak position and double of minimum peak position plus arbitrary positive small integer, it regards the PM as the first harmonic position and takes half of the PM. Otherwise it takes PM itself as a fundamental frequency.

Harmonic Product Spectrum (HPS) [2]: The harmonic product spectrum is popular frequency domain algorithm. It finds the peaks from the spectrum of signal using the fact that harmonics have integer overtone relation with the base frequency. It resamples the signal spectrum Y at integer times the original sampling rate, and product-resampled spectrums, Y_m. If there is a base frequency, it also has harmonics which frequency is multiple of base frequency. A frequency which magnitude is max of Y_m is base frequency. We use only original signal and 2 times over-sampling signal to find base frequency due to narrow frequency domain filtered band-pass filter.

Simple Autocorrelation [5]: The simple autocorrelation algorithm use autocorrelation function to get fundamental frequency from PPG signal, which MA are removed and band passed by low cutoff frequency 0.4Hz and high cutoff frequency 0.5. After autocorrelation function it goes through quadratic interpolation with three adjacent samples. Finally it returns first peak of the result from autocorrelation, which is interpolated by quadratic function.

Revised YIN [3]: This algorithm is based on original YIN [3], but it has some differences in some portions such as removing harmonic signals, usage of aperiodicity characteristics, peak selection and final verification. First, it forms the autocorrelation function $r(\tau)$ and squared difference function $d(\tau)$ as follows,

$$r(\tau) = \sum_{i=1}^{W} x_i x_{i+\tau}$$

$$d(\tau) = \sum_{i=1}^{W} (x_i - x_{i+\tau})^2$$

where x is ppg signals after SSA, W is a value that subtract maximum period τ_{max} from the length of x and τ is several period candidates which range is 25 to 313. From $r(\tau)$ and $d(\tau)$ it derives tapered autocorrelation function $r_2(\tau)$ cumulative mean normalized difference function $d'(\tau)$.

$$r_2(\tau) = \frac{r(\tau)}{\left(1 - \frac{\tau - \tau_{min}}{\tau_{max} - \tau_{min}}\right)}$$

$$d'(\tau) = \frac{d(\tau)}{\frac{1}{\tau}\sum_{i=1}^{W} d(\tau)}$$

After that it refines the local minima of $d'(\tau)$ with parabolic interpolation of all triplets. Then, with the common value of local maxima of $r_2(\tau)$ and local minima of $d'(\tau)$, it tries to find the closest to the maximum value of $d'(\tau)$ and let those name as the prd. Among them the value which aperiodicity is smallest is the initial period candidate. Aperiodicity is derived as follow.

$$ap(prd) = \frac{\sum_{i=1}^{W}(x_i^2 + x_{i+prd}^2)}{\sum_{i=1}^{W}\left(x_i - x_{i+prd}\right)^2}$$

Then finally we can get the last estimated period value through the refined initial period candidate by parabolic interpolation.

Nssa: Nssa is obtained from SSA procedure. Let the PPG signal after SSA process is Y_B. Nssa is the maximum peak of the Y_B spectrum.

[1] 10 is the index correspond to the tolerance frequency of HR 0.3Hz in SSA.

YBLS: Y_{BLS} is also obtained from SSA process. However Y_{BLS} is based on another find-peaks function. We use new function named 'findpeaksLSI'. It find peaks in data using clustering algorithm, therefore it can distinct noise's peaks and 'real' peak. Y_{BLS} is a result of two option. First option is to detect harmonics. If it detect that Y_{BLS} is overtone of some frequency Y_{BASE}, select Y_{BASE}. Second option is to detect whether this is correct direction or not. If it points to far away frequency index more than 10 times, then goes there.

Global Maximum (GM): The Global Maximum (GM) use the reconstructed data after SSR and finds the maximum index value inside the range of frequency, from low cutoff frequency 0.4Hz to high cutoff frequency 5Hz.

2) *Voting:* In the voting stage, SPT uses seven pitch-tracking algorithms to make result from the raw data. seven algorithms are as follows, Spectral Peak Selection, Harmonic Product Spectrum (HPS), Simple Autocorrelation, YIN, Nssa, YBLS and Global Maximum (GM). The voting system makes seven candidates from the seven algorithms and generates the weighted votes. The voting system calculates the distance between the candidates and others to figure out whether the distance is below the offset parameter. Before determine the distances, it products high positive integer if the candidate is close to the first and second maximum MA location. After that if the candidates are in the range of offset parameter the voting system will add the vote value of the candidate based on the weight parameter. Each candidate has its own weight parameters.

The weight parameters are formed based on the error rate. The equation is as follows

$$weight = \frac{1}{Avg * Std}$$

where Avg stands for the average of error rate and Std stands for the standard deviation of error rate. In the proposed algorithm, the weight parameters are shown on the Table I. The proposed weight parameters are calculated by testing 30 times.

3) *Verification:* In the verification stage, it makes a final output value based on the previous estimated data. In the first time-window, the proposed algorithm will use the value, selected from the voting system because there is no previous estimated data on the first time-window. After first time-window, there are two methods to verify the result.

The first method is using the result of voting system directly. This method is used when the number of candidates that are located inside the range of offset parameter exceeds the number of 9. It is a serious problem if the previous values are on the wrong location. By using this method, the proposed algorithm will go back to the correct location faster than not using it. The number of 9 will guarantee that the selected candidate is the right answer.

Second method is using tau parameter. If the difference between selected value from the voting system and the previous value exceeds the offset parameter, the verification system will add tau parameter to the previous value, and if the difference between selected value from the voting system and the previous value is smaller than the negative number of offset parameter, the system will subtract tau parameter to the previous value. The calculated value will be the final result of the current time window. The equation is as follows

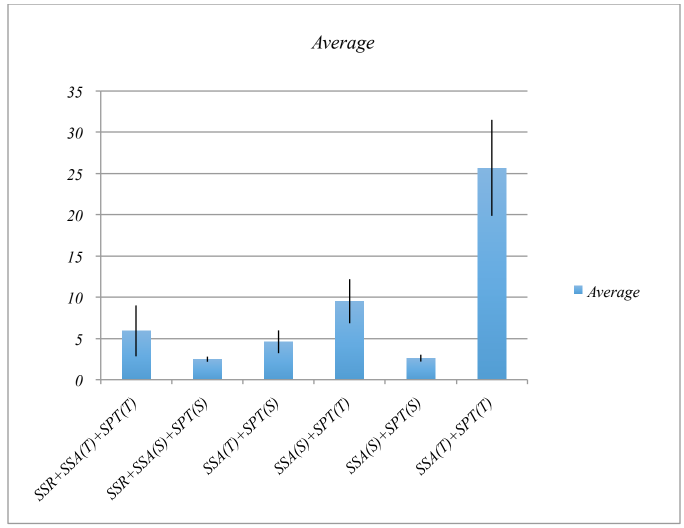

Fig. 1. Compare SSA and SPT of TROIKA and STORM framework by comparing each average error.

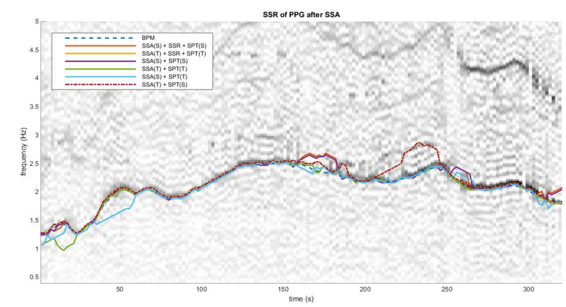

Fig 2. Comparing the result of subject 8

$$N_{\text{curr}} = \begin{cases} N_{\text{prev}} + \tau, & if\ N_{\text{voting}} - N_{\text{prev}} > \quad offset \\ N_{\text{prev}} - \tau, & if\ N_{\text{voting}} - N_{\text{prev}} < -offset \\ N_{\text{voting}} & \end{cases}$$

In our algorithm, tau parameter is set to 2 and offset parameter is set to 6.

III. Experimental Results

We used 12 training sets, which are provided on the SPCUP site: (http://www.zhilinzhang.com/spcup2015/data.html). We also made an imperfect implementation of SSA and SPT based on the TROIKA [1], and tested it with those 12 training sets. Fig. 1 above is a comparison of SSA and SPT between TROIKA and our proposed algorithm STORM (T) stands for the TROIKA framework, whereas (S) stands for the STORM. As presented on the Fig. 1, using STORM SSA and SPT makes the most significant result. Also, the combination including SPT(S), which is the key part of the proposed algorithm, makes the lowest error rate

Fig. 2 shows the result graph of the tested set. This result include SSR+SSA(S)+SPT(S), SSR+SSA(T)+SPT(T), SSA(S)+SPT(T), SSA(T)+SPT(S), SSA(S)+SPT(S), SSA(T)+SPT(T) and the answer BPM value.

Because TROIKA framework is not disclosed, we had to make an imperfect implementation. There, it can be little different from the original TROIKA algorithm.

IV. Conclusion

In this paper, we proposed a general algorithm called STORM to find out the most accurate estimation of PPG signals. The proposed algorithm consists of three parts, Singular Spectrum Analysis (SSA), Sparse Signal Reconstruction (SSR) and Spectral Peak Tracking (SPT). The most important part of this algorithm is SPT, which include voting system. The proposed algorithm uses 7 algorithms to estimate the data and this makes the algorithm more general than any other PPG estimating algorithms. Each part is necessary to generate the proposed algorithm. By this general and powerful algorithm, the wearable devices can get the most accurate information of heart rate from the PPG signals

References

[1] Z. Zhang, Z. Pi, B. Liu, "TROIKA: a general framework for heart rate monitoring using wrist-type photoplethysmographic (PPG) signals during intensive physical exercise", *IEEE Transactions on Biomedical Engineering, in press*, 2014

[2] A.M.Noll. "Pitch determination of human speech by the harmonic product spectrum, the harmonic sum spectrum, and a maximum likelihood estimate", *volume 19 of Symposium on Computer Processing in Communication, pages 779-797. Polytechnic Institute of Brooklyn Microwave Research Institute, New York*, 1970.

[3] A.D.Cheveigne, H.Kawahara, "YIN, a fundamental frequency estimator for speech and music", *Acoustical Society of America, in press,* 2002.

[4] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm," *IEEE Trans. on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.

[5] Armani, L and Omologo, M, "Weighted autocorrelation-based F0 estimation for distant-talking interaction with a distributed microphone network", *IEEE International Conference, Acoustics, Speech, and Signal Processing, 2004*