A New Method for Instantaneous Signal Period Identification by Repetitive Pattern Matching

Michael Liscombe and Amir Asif

Department of Computer Science and Engineering

York University, Toronto, ON, Canada M3J 1P3

Abstract—**Identifying reoccurring patterns within a signal set often yields important information about the signal. Fourier analysis is typically used as a method to determine periodic sinusoidal elements of a given signal; however, it is not able to properly identify near-periodic patterns with small instantaneous changes in the periodicity of the input waveform. To that end, a new transform has been developed which performs a form of high gain correlation on a selected subset of time-localized data points. This transform yields a periodicity vector which accurately identifies a set of samples which are found to be approximately repetitious over a two period timeframe. The transform has been extended to the problem of heart beat detection, referred to as QRS Complex detection in the literature.**

I. Introduction

The problem of signal period identification has been studied for many years and has seen many new analysis techniques created. Fourier analysis has been accepted as the standard for determining the periodicity of a signal, but has been shown to have limitations. Other transforms such as the Short-Time Fourier Transform (STFT) [1] and Stockwell Transform [2] have been created to tackle the problem of small aperiodic signal variations; however, these time-frequency transforms are still limited to the same constructs as the Fourier transform, namely, the translation of a signal into periodic sine waves. As such, performing time-frequency analysis on a near periodic, repetitious pattern in the time domain result in an aperiodic repetitious pattern along the time axis in the time-frequency domain. However, such a result requires further analysis in 2D space to acquire the period of the original repetitious pattern; thus, the problem of identifying the instantaneous period of a repetitious, aperiodic signal pattern cannot be solved using standard frequency or time-frequency analysis tools.

A new algorithm has been created based on the Modified P-Spectrum approach. The original P-Spectrum transform can be used to determine the periodicity of an arbitrary, repetitious pattern in a given waveform. It was later modified by Qui and Liu [3] to allow for identification of aperiodic, repetitious patterns, resulting in the creation of the Modified P-Spectrum. The algorithm presented in this paper proposes a new method of calculating the Modified P-Spectrum, which provides significant reductions in both computational complexity and processing time. Henceforth, this innovative method will be referred to as the Enhanced Modified P-Spectrum transform.
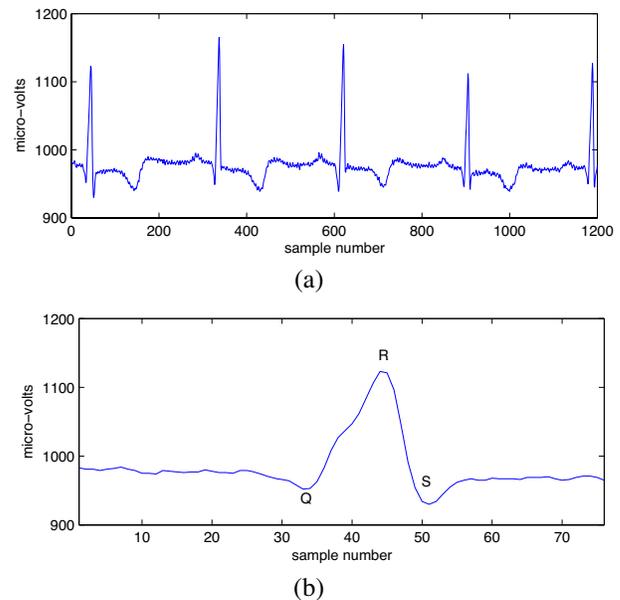
Fig. 1. Electrocardiogram (ECG) of the electrical activity in the heart: (a) Original, and; (b) Expanded to show the QRS Complex.

To better exemplify the use of the proposed algorithm, this paper will be based on the problem of heart beat detection, referred to as QRS Complex detection in the literature.

The electrocardiogram (ECG) is a non-invasive measurement of the electrical activity in the heart (Fig. 1(a)). When conducting a 12-lead ECG, 12 electrodes are placed on selected locations on the patient's body and the voltage difference between the electrodes is recorded. Analysis of the resulting ECG waveforms yields important information for diagnosing the well-being of the heart and also helps in identifying weaknesses in the heart muscles. A heart beat consists of {Q, R, S}-waves and, as such, is often referred to as the QRS complex. The amplitude and duration of these waves is often studied, since these parameters determine the shape of the QRS complex (Fig. 1(b)). The heart beat rate, or heart-rate in short, is calculated by determining the interval between two R-waves and solving for beats-per-minute (bpm). While collecting heart-rate data, a record of all R-R intervals (also referred to as the R-R series) is formed. This data series is commonly used in the diagnosis of tachycardia, bradycardia, and arrhythmia by performing a study of the variation in the beat-to-beat intervals, known in the literature as heart rate

variability (HRV). Clearly, automatic analysis of ECG parameters such as the $\{Q, R, S\}$-waves, heart-rate, and R-R interval first require the detection of the QRS complex. The problem of QRS complex detection has been studied extensively and many high performance algorithms have been proposed [5]. Although several methods yield excellent detection results, they also require fine-tuned thresholds, training data, or other a priori knowledge; therefore, most of these algorithms are limited to offline detection, or in many cases have high processing requirements which preclude them from real-time applications. Computationally efficient run-time performance is desirable for extending the battery life of an instrument such as a mobile Holter recording device. Another modern application includes the incorporation of ECG software into mobile phones [6], an adaptation that benefits greatly from a more computationally efficient QRS complex detection algorithm.

The paper presents a computationally efficient implementation of the modified p-spectrum transform [3] for detecting the QRS complex in ECG waveforms. The proposed algorithm is unaffected by the beat-to-beat scaling of the QRS complex or changes in the R-R interval. As well, detection does not depend on the shape or characteristics of the QRS complex within a given period; only on the similarity between two consecutive periods of the ECG waveform. The paper is organized into four sections. Section II describes four algorithms for QRS detection: (i) The p-spectrum; (ii) The modified p-spectrum proposed in [3] that enhances the runtime speed of the p-spectrum approach; (iii) The qrs-spectrum that combines the modified p-spectrum with a peak detection approach; (iv) the enhanced modified p-spectrum, and; (v) the enhanced qrs-spectrum. Schemes (iv) and (v) are new approaches proposed in this paper. Section III illustrates the speed-up obtained with the proposed approaches. Section IV concludes the paper.

II. METHOD

A. The P-Spectrum

The p-spectrum is a tool used to determine if a signal is periodic. Given a discrete-time data series $X = [x_1, x_2, \ldots, x_N]$, the p-spectrum can be calculated using the following three steps. For all values of p, $(2 \leq p \leq \lfloor N/2 \rfloor)$,

1) Create the matrix A_p defined as

$$A_p = \begin{bmatrix} x_1 & x_2 & \ldots & x_p \\ x_{p+1} & x_{p+2} & \ldots & x_{2p} \\ . & . & \ddots & . \\ x_{mp+1} & x_{mp+2} & \ldots & x_N \end{bmatrix}. \quad (1)$$

If p is not a multiple of N, the vacant entries in the last row of A_p are padded with zero values.
2) Perform the singular value decomposition on A_p.
3) Determine the ratio of the largest and smallest eigenvalues of A_p. The resulting value is the amplitude at period p of the p-spectrum, denoted by p_spec(p).

A signal X is said to be periodic with a fundamental period of p_0, if a relatively large value or "spike" occurs at $p = p_0$ in the p-spectrum. As expected, relatively large values or peaks may
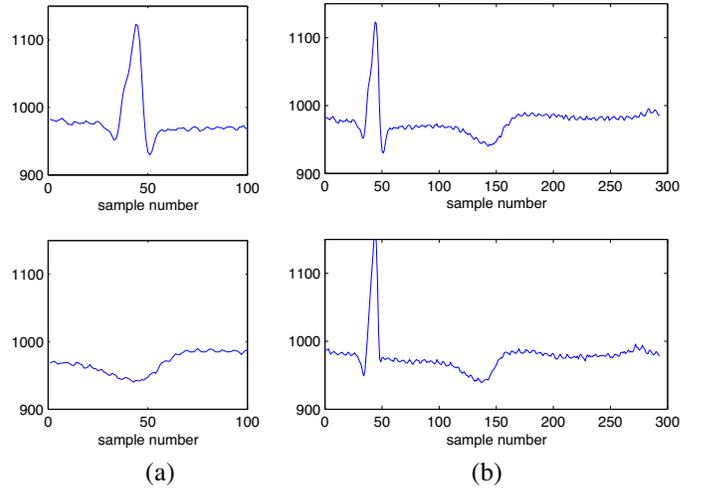

(a) (b)

Fig. 2. Examples for matrix B_p for: (a) $p = 100$, and; (b) $p = 193$. For each case, the top waveform plots the first row in B_p, while the bottom waveform plots the second row in B_p. The y-axis represents the recorded voltage for the respective row in μV.

also be observed due to higher order harmonics at multiples $(2p, 3p,$ and so on) of p.

As a general tool, the p-spectrum provides accurate determination of a truly periodic signal. However, complications arise when ECG datasets are analyzed using this approach. Due to inherent imperfections of the human anatomy, heart beats do not occur at precise intervals. Instead, the heart-rate is subject to constant variations known as heart rate variability (HRV). When the matrix A_p is created, not all of the heartbeats line-up and the ratio of the eigenvalues gets reduced. Typically, the p-spectrum yields a relatively low amplitude at period p_0.

B. The Modified P-Spectrum

The modified p-spectrum [3] is an altered version of the original p-spectrum algorithm designed specifically for signals that have variable periodicity. The transform is determined using the same steps as the original p-spectrum approach, except that A_p is replaced by the rectangular $(2 \times p)$ matrix

$$B_p = \begin{bmatrix} B1_p \\ B2_p \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \ldots & x_p \\ x_{p+1} & x_{p+2} & \ldots & x_{2p} \end{bmatrix}. \quad (2)$$

The ratio between the maximum and minimum eigenvalues obtained from the modified p-spectrum for a given value of p is denoted by mp_spec(p). Fig. 2 plots the first row, $B1_p$, and second row, $B2_p$, derived from an ECG signal for $p = 100$ and $p = 293$. In Fig. 2(a), the value of p is set to 100 and the two rows of B_p are dissimilar resulting in a small value for mp_spec(p). Contrarily, Fig. 2(b) plots the two rows in B_p for $p = 293$. The two rows are very similar resulting in a large value for mp_spec(p). The resulting p-spectrum obtained by plotting mp_spec(p) as a function of p is shown in Fig. 3(a). *Runtime Complexity of the Modified P-spectrum:* In the following analysis, the ECG waveform is divided into smaller blocks of length N which are then used in the modified p-spectrum for QRS complex detection. The major computa-

tional step in the modified p-spectrum algorithm is to perform SVD on B_p. For an $(n \times p)$ matrix, where $(n \leq p)$, the deterministic SVD operation has a computational complexity of $O(pn^2 + p^2 n)$, [7]. In our algorithm, $n = 2$, so the SVD runtime complexity is $O(4p + 2p^2) = O(p^2)$ for a fixed value of p. Since p varies from 2 to $\lfloor N/2 \rfloor$, the order of computations is

$$\sum_{p=2}^{\lfloor N/2 \rfloor} p^2 \approx \frac{N(N+1)(N+2)}{24}, \text{ or, simply} O(N^3). \quad (3)$$

C. The qrs-Spectrum

To improve QRS complex detection, another spectrum (auxiliary spectrum), given below, is created to complement the detection capability of the modified p-spectrum.

For all values of p, $(2 \leq p \leq \lfloor N/2 \rfloor)$, compute

$$\text{aux}(p) = \max \left(\frac{B1_p \oplus B2_p}{2} \right), \quad (4)$$

where $\oplus$ is the row-wise addition operator. For each value of p, a row-wise average of $B1_p$ and $B2_p$ is performed, forming a $(1 \times p)$ vector. The maximum value in this vector is determined and stored in position p in the auxiliary spectrum. Referring back to Fig. 2(a) with $p = 100$, after a row-wise average is performed, the maximum value in aux(100) is found to be relatively low since the peaks do not overlap. Performing the same calculation, for values of B_p shown in Fig. 2(b) with $p = 293$, yields a substantially higher maximum value due to the overlapping peaks. The final plot for aux(p) as a function of p is shown in Fig. 3(b).

In-place multiplication of the modified p-spectrum and the auxiliary spectrum defined as

$$\text{qrs\_spec}(p) = \text{mp\_spec}(p) \otimes \text{aux}(p) \quad (5)$$

is used to form the qrs-spectrum. In Eq. (5), symbol $\otimes$ is the row-wise multiplication operator. The qrs-spectrum combines two forms of QRS complex detection. The first method, the modified p-spectrum approach, uses the shape of two consecutive periods of a waveform to locate the best scale-independent pattern match. The second method uses a simple peak detection algorithm, which tries to determine the most likely position of a peak for a given value of p. Intuitively, the combination of shape and peak detection (Eq. (5)) should yield a stronger detection algorithm. Fig. 3(c) uses Eq. (5) to combine the results plotted in Figs. 3(a) and 3(b).

Runtime Complexity of the QRS Spectrum: The auxiliary spectrum computes the in-place addition of two vectors of length p, divides the results by two, and then finds the maximum value in the resulting dataset. These operations are performed sequentially and are each of $O(p)$. As in the modified p-spectrum, p varies from 2 to $\lfloor N/2 \rfloor$ in the qrs-spectrum. The order of computations is

$$\sum_{p=2}^{\lfloor N/2 \rfloor} p \approx \frac{(N-2)(N+6)}{8}, \text{ or, simply} O(N^2). \quad (6)$$

Since the qrs-spectrum uses both the auxiliary and modified p-spectra (Eq. 5), the overall complexity is still of $O(N^3)$.
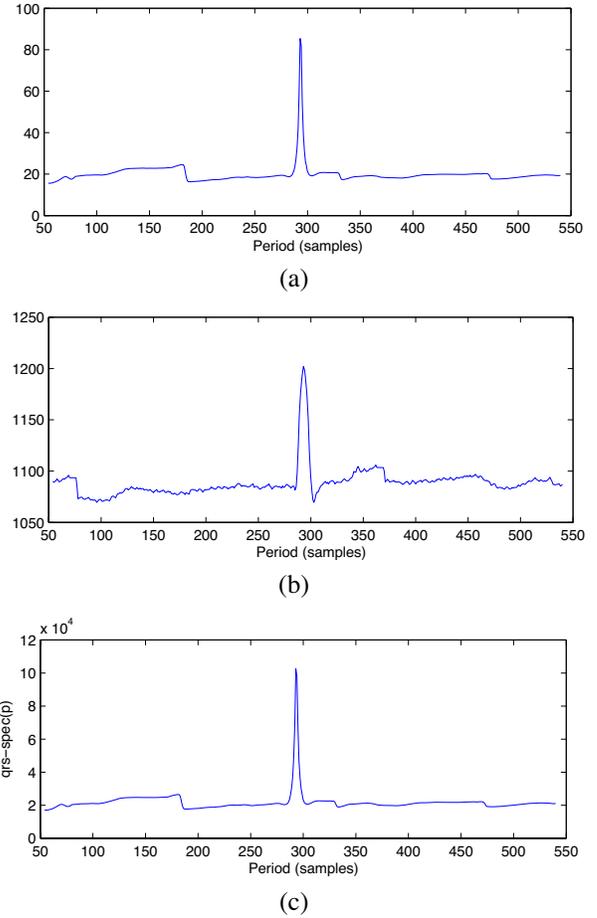

Fig. 3. The resulting p-spectra obtained from: (a) the modified p-spectrum approach; (b) the auxiliary spectrum approach (Eq. (4)) and; (c) the qrs-spectrum approach (Eq. (5)). Within a scaling factor, the result for the enhanced modified p-spectrum is almost the same as that for the modified p-spectrum (plot (a)) and, therefore, is not included. Likewise, the results for the enhanced qrs-spectrum and original qrs-spectrum (plot (c)) are the same within a scaling factor.

D. The Enhanced Modified P-Spectrum

In the modified p-spectrum algorithm, SVD is computationally intense. We propose a new method of vector comparison that requires substantially less computation time. Steps 2 and 3 of the modified p-spectrum algorithm are changed as follows:

2) Compute

$$\theta_p = \cos^{-1} \left(\frac{< B1_p, B2_p >}{||B1_p|| ||B2_p||} \right), \quad (7)$$

where $< \cdot >$ denotes the dot product and $|| \cdot ||$ is the L2 norm, or magnitude, of the enclosed vector.

3) Using the value of θ_p obtained in Step 2, compute the enhanced modified p-spectrum at p as

$$\text{emp\_spec}(p) = \frac{1}{\theta_p}. \quad (8)$$

The enhanced modified p-spectrum, proposed in this paper, can be envisioned by considering two arbitrary vectors of length p in a p-dimensional space. In our approach, the angle

between these vectors is used as a measure of their similarity. The smaller the angle, higher is the similarity between the two vectors. For example, if the two vectors are parallel, they only differ by a multiplicative constant and the angle between them is zero. We use the inverse of this angle (denoted by emp_spec(p)) as a measure of similarity. Larger the inverse, greater is the similarity, and vice versa.

Runtime Complexity of the Enhanced Modified P-Spectrum: The runtime of the enhanced modified p-spectrum is found to be $O(p)$, as shown below.

$$\left\{ <B1_p, B2_p> = \sum_{i=1}^{p} B1_p(i)B2_p(i) \right\} \Longrightarrow O(p) \quad (9)$$

$$\text{and } \left\{ ||Bj_p|| = \sqrt{\sum_{i=1}^{p} |Bj_p(i)|^2} \right\} \Longrightarrow O(p), \quad (10)$$

for $j = 1, 2$. In the above equations, $B1_p(i)$ refers to the i'th element in $B1_p$, and similarly for $B2_p(i)$. The above calculations are performed sequentially; thus, the total runtime is found to be $O(p)$, with the arccos lookup table and division being computed in $O(1)$ time. As shown in Eq. (6), the overall computational complexity for an ECG of length N samples is of $O(N^2)$ for the enhanced modified p-spectrum.

E. The Enhanced QRS- Spectrum

The enhanced modified p-spectrum (Section II-D) is found to yield very similar results when compared with the original modified p-spectrum (Section II-B). Following the steps outlined in Section II-C, an improved implementation of the qrs-spectrum is obtained by replacing the modified p-spectrum with its enhanced version as follows

$$\text{eqrs\_spec}(p) = \text{emp\_spec}(p) \otimes \text{aux}(p) \quad (11)$$

where emp_spec(p) is the proposed approach presented earlier. *Runtime Complexity for The Enhanced QRS-Spectrum:* In Section II-C, the auxiliary spectrum was shown to run in $O(p)$ time. As well, the enhanced modified p-spectrum was found to have a runtime complexity of $O(p)$ in Section II-D. The enhanced qrs-spectrum is formed by the in-place multiplication of the auxiliary spectrum and enhanced modified p-spectrum, an $O(p)$ operation. The total runtime complexity of the enhanced qrs-spectrum is, therefore, of $O(p)$ for a fixed value of p. As proved before, the overall complexity of the enhanced qrs-spectrum is given by $O(N^2)$ for an ECG block of length N samples.

As a final note to our discussion, we observe that for ECG signals, the range p of the spectrum is limited to the upper and lower bounds

$$p_{\min} = \left\lfloor \frac{60 f_s}{HR_{\max}} \right\rfloor \qquad p_{\max} = \left\lfloor \frac{60 f_s}{HR_{\min}} \right\rfloor \quad (12)$$

where f_s is the sampling rate used to digitize the ECG and $HR_{\min}$ ($HR_{\max}$) is the minimum (maximum) heart-rate to be detected. Based on a sampling rate of 360Hz with $HR_{\min}$ and $HR_{\max}$ set to 40bpm and 400bpm, respectively, the minimum and maximum values of p used in the study are 54 and 540.

III. RESULTS

The results section compares the performance of the original qrs-spectrum (Algorithm C) with that of the enhanced qrs-spectrum (Algorithm E). The goal of this section is to show that the enhanced qrs-spectrum transform is faster and that it yields results that are approximately equivalent to those of the original qrs-spectrum; therefore, both methods should yield roughly the same performance characteristics when performing QRS complex detection on an ECG. While determining the runtime of the qrs-spectrum, we use the thin SVD implementation [7] for computing the eigenvalue ratio. For a $(n \times p)$ matrix of rank r, the computational complexity of the thin SVD implementation is $O(npr^2)$, which is a factor p less than the general implementation of SVD.

a) Computational Complexity: Runtime Comparison: The original and enhanced qrs-spectrum algorithms differ by only two steps (Steps 2 and 3 as described in Section II-D), therefore, it was sufficient to determine the difference in computation times for these two steps alone. From this point onwards, the runtimes described for the two algorithms refer only to the two differing steps in each transform. The run-times reported in the paper are based on experiments performed using a 2.4GHz desktop computer with 1GB RAM. A wide array of sampling frequencies from $f_{s(\min)}$ = 10Hz to $f_{s(\max)}$ = 10KHz were used in our experiments. The following analysis was performed using $HR_{\min}$ = 40bpm and $HR_{\max}$ = 400bpm. For each value of the sampling frequency f_s within the range ($f_{s(\min)} \leq f_s \leq f_{s(\max)}$) and an increment of 1Hz, the following steps are used to determine the computation time for the two algorithms as a function of f_s.

1) Compute $p_{\min}$ and $p_{\max}$ based on Eq. (12).
2) Create vectors t_{svd} and t_θ, which contain, respectively, the runtimes of the original and enhanced qrs-spectrum as a function of the sampling frequency f_s.
3) For each value of p such that ($p_{\min} \leq p \leq p_{\max}$), do
   (a) Generate and save the B_p matrix.
   (b) Compute qrs_spec(p) by evaluating the ratio of the largest and smallest eigenvalues of B_p.
   (c) Compute eqrs_spec(p) by taking the inverse of θ_p, defined in Eq. (7).
4) Sum the computation times across p to obtain the runtime of the original and enhanced qrs-spectrum algorithms. Store the results in t_{svd} and t_θ as a function of the sampling frequency f_s.

The runtime vectors t_{svd} and t_θ are averaged across different ECG signals based on a Monte-Carlo simulation. The average runtimes for the original and enhanced qrs-spectrum algorithms are plotted in Fig. 4, which shows that the enhanced qrs-spectrum approach requires substantially less runtime than the original method. Fig. 5 illustrates the achievable speedup with the enhanced approach by plotting the relative difference ($[t_{svd} - t_\theta]/t_\theta$) of the two runtimes. It is clear that the enhanced qrs-algorithm provides a speedup of a factor of 3 to 4 over the original qrs-algorithm.
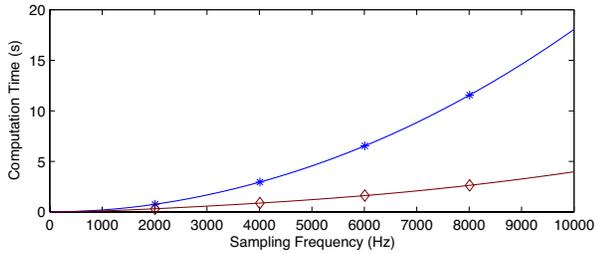
Fig. 4.   Runtime for: (a) the original qrs- (plotted as '∗'), and; (b) the enhanced qrs-spectrum (plotted as '◇').
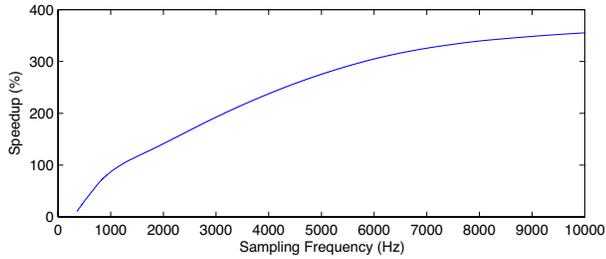

Fig. 6.   Maximum allowable sampling rate for heat-rates constrained to 400bpm for the original qrs- ('∗') vs. enhanced qrs- ('◇') spetrum.

Fig. 5.   Speedup with the enhanced qrs-spectrum algorithm versus the original qrs-spectrum algorithm.

Fig. 7.   Ratio(p) of the actual outputs of the original and enhanced qrs-spectrum algorithms as a function of p.

The authors of the original qrs-spectrum algorithm designed their detection system to function with heart-rates of up to 400bpm, corresponding to a period of 0.15s. For comparative purposes, we have imposed the same constraint for our detection system in Fig. 6, where the period of 0.15s is plotted as a dotted line. Assuming real-time computations with the results being displayed as the ECG is being recorded, Fig. 6 shows that the absolute maximum sampling frequency that can be used to perform detection for the original qrs-spectrum algorithm is 775Hz. For the enhanced approach, the sampling frequency is 1225Hz (an improvement of 58%), thus, allowing for a higher resolution in our estimation.

b) Qualitative Analysis: A first order analysis of the original and enhanced qrs-spectrum algorithms yielded an interesting observation: the spectra were multiples of each other. To test this hypothesis, Fig. 7 plots the ratio

$$\text{ratio}(p) = \text{mp\_spec}(p)/\text{emp\_spec}(p). \qquad (13)$$

as a function of p. The above ratio is equivalent to

$$\text{ratio}(p) = \text{qrs\_spec}(p)/\text{emp\_spec}(p), \qquad (14)$$

the ratio of the actual outputs of the original and enhanced qrs-spectrum algorithms. Fig. 7 illustrates that the two algorithms produce the same result with a difference of a multiplicative factor of 2. Averaging over p, the spectra in Fig. 7 has a mean of 1.9991 and a standard deviation of 1.1930×10^{-4}. Note that the standard deviation is very small, meaning that the majority of these values are close to the mean. These minor variations are attributed to the numerical precision of the system.

IV. CONCLUSION

A new algorithm that determines the instantaneous period of a near periodic, repetitive pattern was presented. We refer to
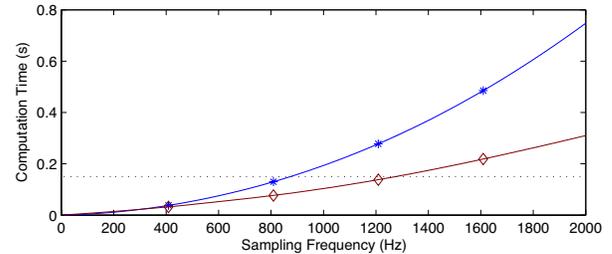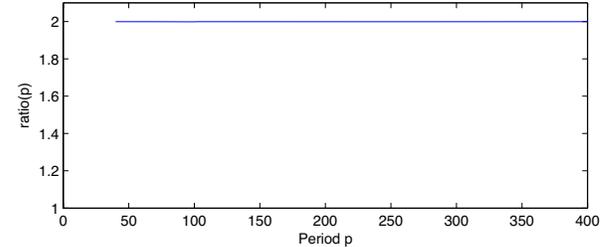
this algorithm as the enhanced qrs-spectrum transform. The algorithm is equivalent to the modified P-spectrum algorithm [3] and produces identical results. The proposed algorithm is also unaffected by the beat-to-beat scaling of the QRS complex or changes in the R-R interval. As in the original approach, detection in the improved implementation does not depend on the shape or characteristics of the QRS complex within a given period; only on the similarity between two consecutive periods of the ECG waveform. The proposed algorithm provides superior runtime performance with a speedup of up to a factor of 3.55 for sampling frequencies ranging from 268Hz to 10KHz over the conventional qrs-spectrum approach. This is fairly remarkable since the conventional approach is optimized for runtime performance and uses the computationally efficient thin SVD algorithm in its implementation.

REFERENCES

[1] M. R. Protnoff, "Time-Frequency Representation of Digital Signals and Systems Based on Short-Time Fourier Analysis," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(1), 1980. pp. 55-69.
[2] R. G. Stockwell, L. Mansinha, and R. P. Lowe., "Localization of the complex spectrum: The S-Transform," IEEE Transactions on Signal Processing, 44(4) 1996. pp. 998-1001.
[3] P. Qui and K. J. R. Liu, "A Robust Method for QRS Detection Based on the Modified P-Spectrum," in Proceedings of ICASSP, 2008.
[4] R. Mark, and G. Moody, MIT-BIH Arrhythmia Database Directory," Cambridge: Massachusetts Institute of Technology, 1988.
[5] B-U. Kohler, C. Hennig, and R. Orglmeister, "The Principles of Software QRS Detection,"*IEEE Eng. in Med. & Biology*, Jan.-Feb. 2002, pp. 42-57.
[6] F. Sufi, Q. Fang, and I. Cosic, "R-R Peak Detection on Mobile Phones," in Proceedings of the Conf. the IEEE Eng. Medicine and Biology Society, 2007, pp. 3697-3700.
[7] G. Golub and C. V. Loan, Matrix Computations, Johns Hopkins University Press, 1996, 2nd edition.
[8] Y. Ishikawa and F. Mochimaru, "Wavelet Theory based Analysis of High frequency, High resolution Electrocardiograms: A New Concept for Clinical Uses," Prog. Biomed. Res., vol. 7, 2002, pp. 179-184.