33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Robust Heart Rate Measurement with Phonocardiogram by On-line Template Extraction and Matching

Yu-Hsin Chen, Hong-Hui Chen, Tung-Chien Chen and Liang-Gee Chen

Abstract— **As health care becomes popular, daily monitoring of health-status related parameters, including the heart rate (HR), is getting more and more valued. An easy, comfortable and robust solution is therefore an important issue. Phonocardiogram (PCG) is a physiological signal reflecting the cardiovascular status. It could be recorded by microphone-equipped on-hand devices, like the smartphone, even without direct skin contact. However, high inter- and intra-variance of PCG make its processing challenging. For PCG-based HR measurement, a robust method is still strongly required. In this paper, we propose a HR measurement algorithm on the processing of PCG that uses on-line template extraction and matching. Through several experiments where traditional methods cannot effectively handle, the robustness of our method is verified by its accurate HR measurement results.**

I. INTRODUCTION

Heart rate (HR) is one of the basic yet important cardiovascular parameters. It has been used from the monitoring of daily living activities to the prediction of acute coronary events [1]. Conventional methods for HR measurement are through clinical electrocardiogram recording devices, pulse oximetry sensors or blood volume pulse sensors. These devices provide a robust means for HR monitoring. Nevertheless, direct skin contact to the device or the adhesive gel patch is required to get accurate records. Moreover, these dedicated hardware might not be easily obtained or accessed, which reduces its practicability in daily preventive actions.

Phonocardiogram (PCG), the recording of heart sounds, is one physiological signal reflecting mechanical heart events. It is commonly used to monitor the status of heart valves and hemodynamics [2]. Its periodic pattern also indicates the possibility for HR measurement. PCG could be recorded using electronic stethoscopes, microphones or on-hand smartphones. With proper consumer electronics equipment, PCG will be easily obtained even without direct skin contact, which points out a good scenario for daily HR monitoring.

The high inter- and intra-individual variance of PCG increases the difficulty on its analysis and processing. Being a non-stationary signal, PCG consists of several different components. The first and second heart sounds (S1, S2) are the most significant ones. There are still other events, nevertheless, that may be recorded, including the third and fourth heart sounds (S3, S4), heart murmur and noise. Fig. 1 demonstrates four PCG recordings with different sound patterns. An unique PCG pattern is anticipated in each individual. Under different activity states, such as at rest or after exercise, PCG patterns also change within each person.

Y.-H. Chen, H.-H. Chen, T.-C. Chen and L.-G. Chen are with the DSP/IC Lab, GIEE, National Taiwan University, Taipei, Taiwan. lgchen@cc.ee.ntu.edu.tw

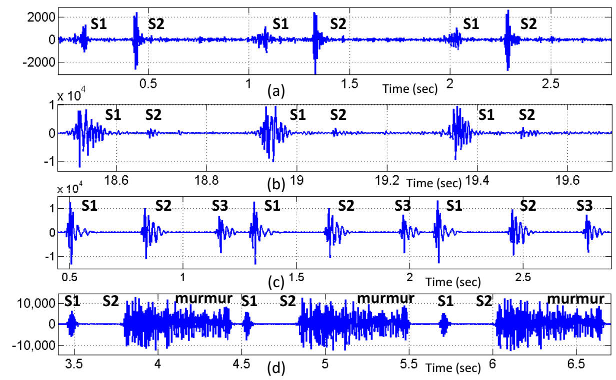

Fig. 1. Several examples of PCG recording show the unique PCG pattern for each individual. (a) and (b) are from the same participant with the former one recorded at rest and the later one recorded after exercise. (c) is the PCG with S3. (d) is recorded from the patient with aortic regurgitation.

The previous works of PCG-based HR measurement usually apply the detection of cardiac pulse peaks [3][4]. These algorithms assume a general heart sound model. With a basic normalization and thresholding on PCG, S1 and S2 are detected, extracted and counted to derive the HR. PCG segmentation techniques that analyze heart sound features are also introduced to make the detection more robust [5][6]. For example, Wavelet Transform is commonly adopted for PCG time-series processing. Mother wavelets like Daubechies family and Morlet are used to intensify S1 and S2. To deal with the signal variance, filtering, spectrum analysis and sophisticated heuristic rules play important roles on the system performance. However, since heart sounds vary in a great extent as shown in Fig. 1, these methods cannot effectively deal with unexpected PCG patterns that differ from the presumed model. For instance, the HR estimation for patients with cardiovascular diseases is most likely to be inaccurate because of the varieties of heart murmur patterns.

In this work, we present an algorithm for the measurement of HR based on PCG. Although a common heart sound model could not be derived, the high autocorrelation of PCG is expected in a short time period. Thus, the concept of on-line template extraction and matching, which is widely used in speech signal processing, provides a promising tool for HR estimation. The influence of inter- and intra-variance of PCG is minimized through this on-line pattern-learning mechanism. With periodically extracted and updated

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

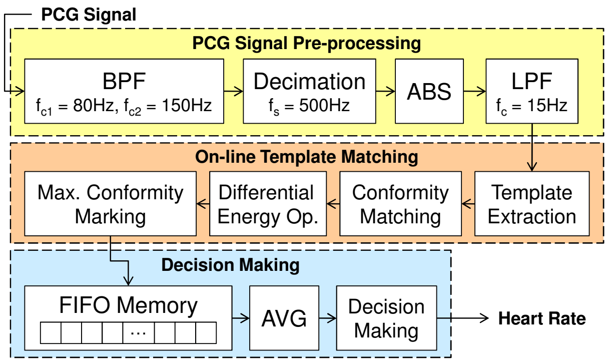

Fig. 2. The block diagram of the proposed algorithm for heart rate measurement.

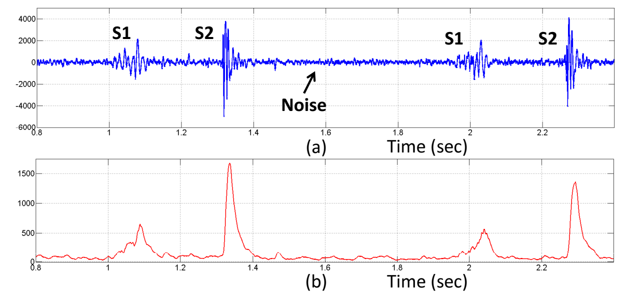

Fig. 3. (a) The original PCG recording and (b) the pre-processed signal envelope. The signal envelope retains the heart sound components S1 and S2 while the noise is suppressed.

templates, repetitive heart sound patterns can be found for an individual, and the heart beat period is then computed. Robust HR measurement is achieved in several scenarios that traditional methods could not function well. This paper is organized as follows. The proposed concept and algorithm details are revealed in Section II. In Section III, we verify the performance of the proposed method through several experiments with the comparison to the traditional methods. Finally, we draw the conclusion in Section IV.

II. The Proposed Method

The detailed block diagram of the proposed algorithm is illustrated in Fig. 2. It could be generally divided into three parts: PCG signal pre-processing, on-line template matching, and decision making. First, the pre-processing stage extracts the signal-of-interest from the original PCG recording. On-line template matching is then performed to find repetitive PCG patterns within the signal. Finally, the decision making stage analyze the matching results to compute the period of the repetitive PCG pattern.

A. PCG Signal Pre-processing

The recording of PCG usually has a sampling frequency higher than 8000Hz. If the recording environment could not be well controlled, noises will be coupled into the PCG. To avoid unpredictable effects brought by noises, filtering becomes important for later processing. Since the main spectrum of S1 and S2 stay within the range of 150Hz, the system filters the original heart sound using a band-pass filter with cutoff frequencies at 80Hz and 150Hz. Then, the signal is down-sampled to the sampling frequency of 500Hz.

In order to efficiently perform time-domain signal template matching, the envelope of the bandpass-filtered signal is approximated to eliminate the influence of phase delay. First, the bandpass-filtered signal is full-wave rectified, and is further filtered using a low-pass filter with cutoff frequency at 15Hz. The effect of the whole pre-processing procedure on the PCG signal is shown in Fig. 3. The envelope of S1 and S2 is clearly identified while the high-frequency background noise is mostly suppressed.

B. On-line Template Matching

Figure 4 demonstrates the procedure of on-line template extraction and matching. The template has the length of N_B samples, and is continuously picked up from the signal envelope with a predefined period T_{update}. For the latest template, conformity matching is performed within a search range. Since the human HR mostly lies in the interval of 45 to 240 beats per minute (BPM), the search range is set between 0.25 and 1.33 seconds later to the template if the test subject is a human being.

In the search range, dozens of signal blocks are compared with the latest template for finding their mutual conformities. The compared signal blocks also have the length of N_B samples. Let the signal envelope samples in time domain be denoted by $\{x_i\}$, the latest template be denoted by Y, and the m$^{\text{th}}$ compared signal block in the search range be denoted by B^m. The conformity c_m between Y and B^m is defined as

$$c_m = \|Y - B^m\| = \sum_{n=1}^{N_B} |x_{j+n} - x_{j+S_m+n}|, \tag{1}$$

where

$$\begin{aligned} Y &= (x_{j+1}, x_{j+2}, \ldots, x_{j+N_B}), \\ B^m &= (x_{j+S_m+1}, x_{j+S_m+2}, \ldots, x_{j+S_m+N_B}), \end{aligned}$$

and S_m is the sample delay between Y and B^m.

If the total number of signal blocks in the search range is K, i.e., m ranges from 1 to K, the conformity values from c_1 to c_K form a series named the conformity trace C. C illustrates the dynamics of template-similarity in the search range. Lower c_m values represent higher probabilities that repetitive PCG pattern appears. Based on this rule, when PCG repeats, the corresponding c_m values will form a valley in trace C. Moreover, if the lowest value in the valley occurs at c_M, the corresponding heart beat period is S_M.

Since the dynamics of the conformity trace is also subject to artifacts and baseline drifts, a transform called Differential Energy Operator (DEO) is introduced, which is defined as

$$d_i = (c_i - c_{i-r}) \times (c_{i+r} - c_i). \tag{2}$$

The DEO transformed trace is $D = (d_1, d_2, \ldots, d_K)$. DEO enhances the part of C with high non-linear energy, such as the valley, while the effects of artifacts and baseline drifts are suppressed. Fig. 5 shows a conformity trace C with its corresponding DEO transformed trace D. In this example, K is 90 and r is 10. The valley is preserved in both traces. The baseline drifts and artifact perturbations are flattened in D.

If the heart beat is fast, one search range will contain more than one repetitive patterns, which results in multiple signal

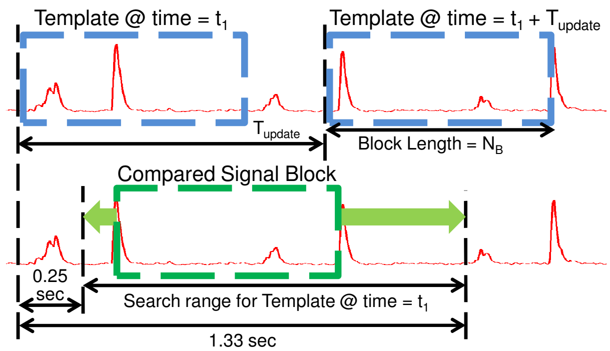

Fig. 4. The upper figure explains the extraction of templates. With the predefined interval T_{update}, templates with length of N_B samples are continuously picked up and updated. The lower figure illustrates the matching procedure. For each template, conformity matching is performed between itself and the signal blocks within the predefined search range.

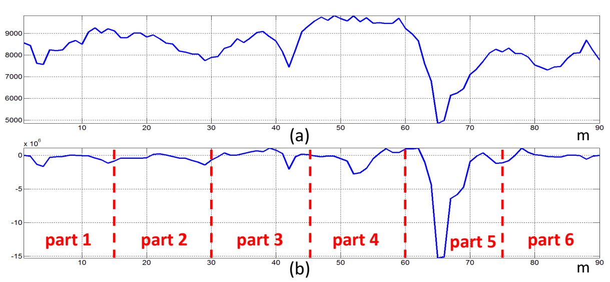

Fig. 5. (a) The conformity trace C and (b) the corresponding trace D after DEO transform. In this example, K is 90 and r is 10. The effects of baseline drift and artifacts are removed after DEO while the valley is more strengthened. For further processing of HR decision making, D trace is divided into 6 parts.

valleys in D. To select the valley that represents the correct heart beat period, a decision making mechanism is required. Thus, the values and locations of these valleys in D should be collected for further processing. D is empirically divided into 6 non-overlapped parts with each part having the same length as shown in Fig. 5(b). In each part, the maximum conformity point (MCP), i.e., the point with the lowest value, is marked, and the information of this 6 MCPs are fed to the next stage for HR decision making.

C. Decision Making

The values and locations of the 6 MCPs in each D are stored in a FIFO fashion. The HR is then computed from the MCPs of consecutive D traces in the FIFO. Let the total number of D traces stored in the FIFO, the value and the location of MCP from part i of nth D trace be denoted as N_D, V_n^i, L_n^i, respectively. The average MCP value AV_i and average MCP location AL_i of part i are defined as

$$AV_i = \textstyle\sum_{n=1}^{N_D} V_n^i / N_D, \tag{3}$$
$$AL_i = \textstyle\sum_{n=1}^{N_D} L_n^i / N_D, \quad i = 1, 2, \ldots 6. \tag{4}$$

AV_1 to AV_6 are then pairwise compared by the decision logic to derive the HR. At each time a new D trace is obtained from the on-line template matching stage as the template is updated, the oldest 6 MCPs in the FIFO are discarded, and the 6 latest MCPs are stored into the FIFO. The average MCP information AV_i and AL_i will be recomputed, and the HR value is ready to be updated. The fastest update period of HR is thus the same with the update period of the template, which is T_{update}.

The decision logic is based on the average MCP information AV_i, AL_i, and the current HR value. If the current heart beat period equals to S_M, the distance between AL_i and M will dynamically determine a threshold Th for AV_i comparison. The rule of the decision logic is

$$\begin{aligned} IF \quad & |AV_i| > |AV_j| \times Th, \ \forall \ j > i \\ THEN \quad & the\ heart\ beat\ period\ is\ S_{AL_i} \\ ELSE\ THEN \quad & i = i + 1, \end{aligned} \tag{5}$$

where $1 \leqslant i, j \leqslant 6$. This rule is initiated from $i = 1$, and is checked until the heart beat period is obtained or $i = 6$. Th is determined as

$$Th = \begin{cases} Th_1, & |AL_i - M| \leqslant K/6 \wedge |AL_j - M| > K/6 \\ Th_2, & |AL_i - M| \leqslant K/6 \wedge |AL_j - M| \leqslant K/6 \quad \vee \\ & |AL_i - M| > K/6 \wedge |AL_j - M| > K/6 \\ Th_3, & |AL_i - M| > K/6 \wedge |AL_j - M| \leqslant K/6 \end{cases} \tag{6}$$

On one hand, since the HR should not change abruptly in a very short period of time, the closer of AL_i (or AL_j) to M means the higher probability that AL_i (or AL_j) is the new heart beat period. On the other hand, if multiple valleys appear, the first valley usually represent the first repeated pattern. Based on these two findings, the relation between the thresholds is $Th_3 \geqslant Th_2 \geqslant Th_1$.

III. RESULTS

A. Experiment Setup

The PCG database with heart murmur patterns is used in the experiment to verify the robustness of our proposed method. The data are acquired from the on-line resources of [7]. We also use the built-in microphone on iPod touch 4th generation by Apple Inc. to record PCG with 16-bit accuracy and 8000Hz sampling frequency. The experiments are made on two males and one female aged from 23 to 32 years old without medical history of cardiovascular diseases. The PCG from two scenarios are considered: (1) the participant is at rest and (2) the participant has just finished an aerobic exercise. The participant wears only one jersey, and the heart sounds is recorded with the microphone placed on the jersey near the mitral and tricuspid area. The recording is performed in an open space with people walking and talking around. Normal breathing is allowed during the experiment.

The parameters of the proposed method is set as follows. The length of the template, N_B, is 500 samples. The update period of the template T_{update} is 0.5 second. This also means that the maximum update rate of HR value is 2Hz. The size of the FIFO, N_D, is set to 10, which implies that one HR value is computed from 5 seconds of heart sound. Based on these settings, the first HR value is obtained around 7 seconds after the PCG recording is started.

The traditional peak detection method is implemented for comparison [4] [6]. It applies the db6 wavelet for wavelet decomposition to intensify S1 and S2. Heuristic rules are also considered to eject redundant peaks. For verification, we manually identify each single heart beat from the PCG waveforms to compute the reference HR.

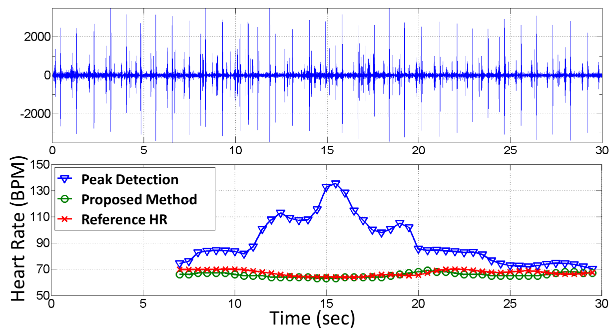

Fig. 6. The upper figure shows the PCG recorded when the participant is at rest. The lower figure shows the comparison of corresponding HR computed from different HR measurement methods.

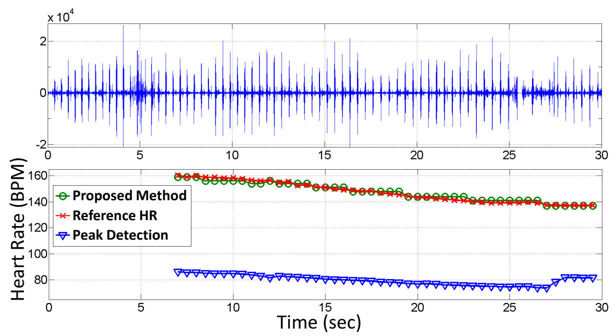

Fig. 7. The PCG recorded when the participant has just finished an aerobic exercise and its corresponding HR results comparison.

B. PCG Recorded at Rest

Figure 6 shows the PCG waveform and the corresponding HR computed from different methods. The PCG is recorded when the participant is at rest. S1 and S2 appear regularly through the waveform. Since the intensity of PCG at rest is relatively weak, the unexpected peaks generated by noise are observed around 10^{th} to 20^{th} second of the recording. The result shows that the peak detection method cannot track the HR when PCG is noisy. The proposed method, however, is less subject to the noise, and its estimated HR is in high accordance with the reference HR. The root mean square error (RMSE) of the peak-detection-based method is 30.15 BPM while the RMSE of the proposed method is 2.40 BPM.

C. PCG Recorded after Aerobic Exercise

Figure 7 demonstrates the PCG waveform and the HR results comparison. The PCG is recorded when the participant has just finished an aerobic exercise. The intensity of S1 is much more larger than S2, which differs a lot from the PCG shown in Fig. 6. In this case, the normalization step of the peak detection method could not effectively characterize S2. The result of the estimated HR by peak detection is then about half of the reference value, since only S1 are extracted. In contrast, the proposed method still catches the dynamics of the HR as it becomes slower with the time. The RMSE of the peak-detection-based method is 67.89 BPM while the RMSE of the proposed method is 1.54 BPM.

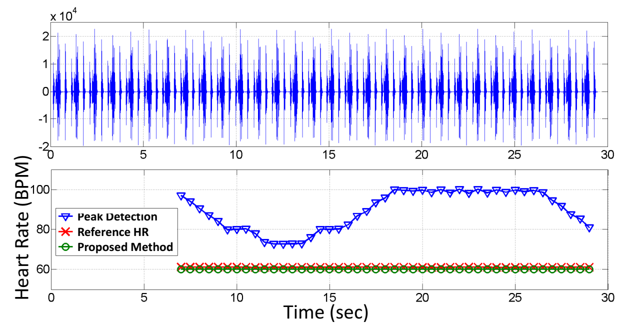

Fig. 8. The PCG recorded from the patient with Mitral Stenosis and its HR results comparison.

D. PCG with Heart Murmur

Figure 8 shows the PCG waveform recorded from the patient with Mitral Stenosis (MS) and the comparison of the HR. This data is from [7]. Due to MS, the murmur patterns are identified in the diastolic period. The peak detection method could hardly differentiate the peaks of S1, S2 and MS, which leads to a wrong peak extraction result. The difficulty on the discrimination of peaks generated by S1, S2 and heart murmurs like MS is also reported in previous works [8]. Thus, the HR estimation based on the peak detection of S1 and S2 is intrinsically impractical. By the proposed method, though, the HR could be accurately tracked even without the pre-knowledge of murmur patterns. For the database of [7], the RMSE of the peak detection method ranges from 0.09 to 34.95 BPM while that of the proposed method lies in the range of 0.02 to 1.08 BPM.

IV. CONCLUSION

The estimation of heart rate based on on-line PCG template extraction and matching turns out to be an effective and robust method. Without prior knowledge on heart sound models, our algorithm dynamically finds the repetitive sound patterns. By this on-line pattern-learning scheme, the influence of inter- and intra-variance of PCG is minimized. It has been shown that the proposed method could report accurate HR values even under the scenarios that traditional peak-detection-based methods cannot effectively handle. These scenarios include noise interference, dramatically changed heart sound intensity, and the appearance of heart murmurs.

REFERENCES

[1] S. Cook and et al., "High heart rate: a cardiovascular risk factor?," *Eur. Heart J.*, vol. 27 no. 20, pp. 2387-2393, 2006.

[2] R. L. Watrous, "Computer-Aided Auscultation of the Heart: From Anatomy and Physiology to Diagnostic Decision Support," *IEEE EMBC*, 2006, pp. 140-143.

[3] M. Godinez and et al., "On-line fetal heart rate monitor by Phonocardiography," *IEEE EMBC*, 2003, pp. 3141-3144.

[4] H. Miwa and K. Sakai, "Development of heart rate and respiration rate measurement system using body-sound," *IEEE ITAB*, 2009, pp. 1-4.

[5] S. Omran and M. Tayel, "A HEART SOUND SEGMENTATION AND FEATURE EXTRACTION ALGORITHM USING WAVELETS," *IEEE ISCCSP*, 2004, pp. 235-238.

[6] Z. Tu and et al., "Improved Methods for Detecting Main Components of Heart Sounds," *IEEE ICNC*, 2010, pp. 3585-3588.

[7] http://solutions.3m.com/wps/portal/3M/en_US/Littmann/stethoscope/education/heart_sounds/

[8] G. Saha and P. Kumar, "An Efficient Heart Sound Segmentation Algorithm for Cardiac Diseases," *IEEE INDICON*, 2004, pp. 344-348.