Multi-Channel Audio-based Estimation of the Pre-Ejection Period

R. P. Paiva, T. Sapata, J. Henriques, I. Quintal, R. Baptista, L. Gonçalves, P. Carvalho

Abstract— **Systolic time intervals (STI) have significant diagnostic and prognostic value to assess the global cardiac function. Presently, STIs are regarded as a promising tool for long-term follow-up of patients with chronic cardiovascular diseases. Heart sound has proven to be a valuable approach for STI estimation, in particular for the Pre-Ejection Period (PEP). However, since the optimal auscultation site varies from individual to individual, as well as with the position of the body, its application in single-channel and fixed auscultation site setups poses practical difficulties. Hence, we extend our previous work on PEP estimation to a multi-channel sound acquisition setup, where signal redundancy is exploited. A channel selection method is proposed and the best channel is selected for PEP estimation. As a preliminary study, the devised algorithms were evaluated with respect to echocardiography reference on a set of 236 heartbeats collected from 8 healthy subjects in two sound auscultation sites. The channel selection approach led to 8.4% estimation error decrease, in comparison to a single-channel approach. Current results support our assumption that a multi-channel audio-based strategy can be applied to assess STI in personal health application scenarios.**

I. Introduction

The timings of myocardial contraction and relaxation are directly related to the health of cardiac cells [1] as they determine the ability of the myocardium to achieve blood delivery according to the metabolic requirements of the organs. The systolic and diastolic timings of the left ventricle assume particular relevance, since this ventricle is responsible for blood flow in the systemic circulation.

Clinically accepted descriptions of systolic function of the left ventricle are the velocity of pressure rise, the velocity of ejection, the extent of ejection and the ejection fraction [2]. These function indicators can be obtained using both invasive and non-invasive procedures (e.g. echocardiography, which is the current gold standard for systolic time intervals measurement) in in-hospital settings. However, these procedures are not adequate for daily applications in home settings, as required for long-term follow-up of patients with chronic cardiovascular diseases. An adequate alternative to evaluate the global cardiac function in this type of application scenarios is the use of systolic time intervals (STI). In fact, several studies have shown that STIs are highly correlated to major and fundamental cardiac functions [3-6].

Systolic ejection is preceded by the electro-mechanical delay and by the isovolumetric contraction time. These two time intervals compose the pre-ejection period (PEP), which is the time interval between the start of ventricular depolarization and the moment of aortic valve opening. The left ventricle ejection time (LVET) is defined as the time interval of left ventricular ejection, which occurs between the opening of the aortic valve and its subsequent closure. PEP and LVET assume major relevance in assessing the cardiac reserve and the left ventricular function [7-8]. PEP is an index of the left ventricular function and reflects changes in myocardial contractility, left ventricular end-diastolic volume and aortic diastolic pressure. Another important application of PEP is in non-invasive beat-by-beat estimation of blood pressure [8]. The left ventricular ejection period can also be related to contractility and to cardiac output [9]. It is by itself a measure of cardiac function.

Several measurement modalities for STI assessment in home settings or other scenarios where portability is recommended have been considered in the literature. Namely, photoplethysmography [10], radial pulse pressure [11], phonocardiography [12] and impedance cardiography [13] are subject of current research efforts.

Heart sound (HS) has proven to be a highly informative diagnostic tool to assess the status of the cardiovascular state of a patient. In fact, in the past we introduced a phonocardiography-based STI measurement algorithm based on a Bayes approach [12], which has shown to exhibit the highest accuracy, precision and correlation with respect to the gold standard among available methods for portable STI measurement [14].

However, the optimal auscultation site varies from individual to individual, as well as with the position of the body. Therefore, single-channel, fixed auscultation site setups, like the one we employed in our past work, pose practical difficulties. On a controlled environment, a single-channel approach could be sufficient in case a specialist would look for and select the best auscultation site. However, this is not feasible for home-monitoring scenarios, where, typically, sound acquisition would be acquired from one or more fixed sites (e.g., sensors integrated into a vest).

To evaluate the potential of this idea, we propose an extension of our previous work on PEP estimation to a multi-channel, fixed sites, sound acquisition setup. Signal redundancy is exploited and the channel with the highest quality is selected, thus allowing for better PEP estimation accuracy. In our experiments, a preliminary two-channel approach achieves 8.4% higher accuracy than the single-channel strategy.

This work was partly supported by the HeartSafe project (PTDC/EEI-PRO/2857/2012) financed by the Portuguese Foundation for Science and Technology and the iCIS project (CENTRO-07-0224-FEDER-002003).

R. P. Paiva, T. Sapata, P. Carvalho and J. Henriques are with CISUC, Department of Informatics Engineering, University of Coimbra, Portugal (e-mail: {ruipedro, sapata, carvalho, jh} @dei.uc.pt).

I. Quintal, R. Baptista and L. Gonçalves are with the Echocardiography Department of the Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal (e-mail: isabelquintal@chc.minsaude.pt).

Preprint submitted to 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Received April 10, 2015.

The remaining sections of the paper are organized as follows. In section II, we describe the experimental setup and data acquisition process. Section III presents the methods for sound channel and PEP estimation. Results are presented and discussed in Section IV. Finally, conclusions from this study are drawn in Section V.

II. Experimental Setup and Data Collection

We carried out a small data collection study involving 8 healthy male volunteers. The data collection was carried out at the Centro Hospitalar de Coimbra (CHC), where heart sounds (two channels) and echocardiographies (echo) were acquired simultaneously. A synchronous ECG with each of the above signals was also acquired and served as a reference signal for co-registration.

The biometric characteristics of the population were:

- Age: 30.4 ± 9.8 years
- BMI: 25.9 ± 4.7 Kg/m2
- Heart rate: 66.0 ± 14.9 bpm

The measurement protocol was conducted by an authorized medical specialist and consisted of several acquisitions of echocardiography in Doppler mode and HS collection in two sites: apex and left sternum border (LSB). More precisely the following steps were carried out:

- The patient was positioned in supine position, turned left (approximately 45º) - the usual echo observation position for the aortic valve;
- The echo was configured for Doppler mode and the stethoscopes were positioned in the LSB and apex regions;
- Runs of 80–100 s. Data acquisitions of HS, echo and ECG were performed repeatedly.

The following signals have been acquired.

- Echocardiography and ECG were acquired using a Vivid system from General Electric. This device produces outputs with images of 500 Hz time resolution.
- HS and ECG: two Meditron stethoscopes and an ADInstruments Bio Amp ECG recorder connected to an ADInstruments Powerlab were applied to record HS and ECG at 2 kHz.

After data acquisition, the annotations of the opening instants of the aortic valve using the echocardiographies were performed under the supervision of an experienced clinical expert in echocardiography. The detected aortic opening times, with the associated PEP values, were used as ground truth for algorithm evaluation.

III. Methods

A. Sound Channel Selection

As previously mentioned, sound was collected from two sites: the apex and left sternum border (LSB).

One solution to select the best channel is to determine the one with the highest signal-to-noise (SNR) ratio. To this end, we propose a signal contrast feature, adapted from the spectral contrast feature proposed in [15].

The method is described in the following paragraphs and is performed for each heartbeat.

Step 1. Select the *audio segment* corresponding to the current heartbeat.

PEP is formally defined as the time interval between the Q-peak of the ECG to the opening of the aortic valve. Hence, an algorithm for detection of Q-peaks [16] is applied. Then, relevant audio segment in each channel starts at the Q-peak time and has a duration of 210 msec. This is motivated by the typical largest PEP durations, as described in [12].

Step 2. Compute the *amplitude contour* of the audio segment corresponding to the current heartbeat, for each sound channel.

The audio segment is full wave rectified and smoothed with a 100 msec half-Hanning window, as in (1):

$$c[n] = |s[n]| * w[n - {}^{W}\!/_{2}], \quad n = 1,2,\cdots,N$$
$$w[n] = 0.5 - 0.5\cos\left(\frac{2\pi(n-1)}{W-1}\right), \quad n = 1,2,\cdots,{}^{W}\!/_{2} \qquad (1)$$

There, $c[n]$ is the amplitude contour of a given audio segment $s[n]$ with N samples, $w[n]$ stands for the half-Hanning window and W denotes the number of samples (corresponding to 100 msec at the employed 2kHz sampling frequency). We guarantee zero output delay by shifting the window.

Step 3. Determine the *signal contrast* of the amplitude contour segment corresponding to the current heartbeat, for each sound channel.

The amplitude contour $c[n]$ is returned as a vector that is then sorted into descending order of magnitude, forming a vector $\{c_1, c_2, \ldots, c_N\}$ such that $c_1 \geq c_2 \geq \ldots \geq c_N$.

The signal peak and valley are then estimated as in (2) and (3), respectively:

$$peak = \frac{1}{\alpha N}\sum_{i=1}^{\alpha N} c_i \qquad (2)$$

$$valley = \frac{1}{\alpha N}\sum_{i=1}^{\alpha N} c_{N-i+1} \qquad (3)$$

There, α is a neighborhood factor set to 0.2 in this article.

Hence, the signal peak is calculated as the average of 20% highest samples in the amplitude contour, while the valley corresponds to the 20% lowest.

Finally, the signal contrast, SC, is computed as the difference between the computed peak and valley, (4):

$$SC = peak - valley \qquad (4)$$

Step 4. Select the *best channel*.

The best channel might be selected in one of two ways:

Preprint submitted to 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Received April 10, 2015.

- On a beat-by-beat mode (BBM), where a different channel might be selected in each heartbeat;
- On complete mode (CM), where the channel that is selected in a higher number of heartbeats is always selected. This allows for inadvertent channel switching when the selection metric selects the wrong channel.

B. PEP Estimation

After selecting the best sound channel (either in beat-by-beat or complete channel mode), an estimation for the PEP duration for each beat is computed.

For the sake of completeness, we summarize our single-channel PEP estimation algorithm described in [12].

As described above, PEP is formally defined as the time interval between the Q-peak (determined as explained above) of the ECG to the opening of the aortic valve. Then, a Bayesian approach is followed for PEP estimation, resorting to the instantaneous amplitude (IA) of the heart sound waveform as the main feature.

First, for each heartbeat, k, the atrio-ventricular (AV) closure time interval, AV_k, is estimated. To this end, the corresponding Q-peak (previously determined) is employed as reference. A Bayesian model is defined, where the prominences, $prom_k$, of the heart sound near the Q-peak, the IA curve, IA_k, and the previous AV interval, AV_{k-1}, are employed, according to (5):

$$\begin{aligned} p(AV_k \mid prom_k, IA_k, AV_{k-1}) \approx \\ \approx p(AV_k \mid prom_k) \cdot p(AV_k \mid IA_k) \cdot \\ \cdot p(AV_k \mid AV_{k-1}) \end{aligned} \quad (5)$$

In brief, the more prominent a peak or valley in the sound waveform is and the higher the IA is, the higher the probability that the AV valves close at that time instant. As for the influence of the previous AV interval, $p(AV_k \mid AV_{k-1})$ is modeled as a Gaussian distribution centered in the previous AV interval, AV_{k-1}, and with a standard deviation $\sigma_{AV} = 30$ *msec*.

PEP duration for heart beat k, PEP_k, is then inferred with recourse to IA_k, AV_k, and the previous PEP interval in beat $k-1$, PEP_{k-1}, as in (6):

$$\begin{aligned} p(PEP_k \mid AV_k, PEP_{k-1}) \approx \\ \approx p(PEP_k \mid AV_k) \cdot p(PEP_k \mid PEP_{k-1}) \end{aligned} \quad (6)$$

In short, the conditional probability distribution of PEP duration given the AV time interval, $p(PEP_k|AV_k)$, is modeled as a Gaussian centered in $AV_k + \mu_{PEP\text{-}AV}$ ($\mu_{PEP\text{-}AV}$ = 30 msec), again with a standard deviation $\sigma_{PEP\text{-}AV} = 30$ msec. This is based on the typical delay between AV closure and aortic valve opening reported by Tavel [17]. In this case, this is modelled as a Gaussian distribution. As for the effect of the previous PEP interval, $p(PEP_k \mid PEP_{k-1})$ is, as before, modeled as a Gaussian distribution centered in the previous PEP interval, PEP_{k-1}, and with a standard deviation $\sigma_{PEP} = 30$ msec.

Additionally, a two-pass approach for PEP estimation is followed. In the first iteration, initial probability distributions (namely Gaussian distributions) for the variables included in the model are assumed. These are parameterized using average population-based values reported in the literature. Then, the initial distributions are updated based on the results obtained. A second run of the algorithm is then conducted using the updated distributions. Thus, the actual model applied is patient-dependent and is estimated using a data-driven approach.

IV. Results and Discussion

Our approach was evaluated in a set of 8 audio clips (one per subject) from the echocardiography-HS collection data, with manually annotated aortic valve openings (based on synchronized echo-cardiography images), comprising 236 heartbeats in total.

Table I summarizes the attained results.

TABLE I
SUMMARY OF RESULTS

Channel	Estimation Error (msec) (average ± std)
Apex	10.12 ± 6.98
LSB	10.13 ± 8.25
Best BBM	9.42 ± 7.12
Best CM	9.27 ± 7.15

As can be seen, the two employed channels, apex and LSB, perform similarly, attaining 10.12 and 10.13 msec average absolute PEP estimation error. After performing channel selection on beat-by-beat mode, the error decreased to 9.42 msec. When channel selection was conducted on complete mode, the estimation error drops to 9.27 msec, i.e., an 8.4% error decrease in comparison to the each individual channel.

In all cases, the standard deviation was around 7 msec, except for LSB, where it raised to 8.25 msec.

Moreover, as illustrated in Table II, the LSB channel was selected much more frequently, indicating that this is a more adequate sound collection site. In fact, in CM mode, the apex was selected for only one out of eight subjects (a particularly thin individual, with BMI = 18.7), amounting for 23 heartbeats. However, this one subject caused the LSB channel to perform slightly worse than the apex channel.

TABLE II
CHANNEL SELECTION RATE

Method	Apex	LSB
Best BBM	16.5%	83.5%
Best CM	9.75%	90.25%

In order to evaluate the significance of our results, statistical significance tests were performed. Namely, the best results, achieved in complete mode, were compared to the absolute errors attained with each channel individually. Since all error distributions were found to be Gaussian using the Kolmogorov–Smirnov test, the paired T-test was carried

out. Comparing CM and LSB results, the results proved statistically significant (p-value < 0.01). However, comparing CM and apex results, a p-value = 0.094 (> 0.01) was found. Hence, in this case the differences did not prove statistically significant. Nevertheless, this is in the border of statistical significance, which could be achieved if a larger dataset had been employed. In any case, the achieved improvement (almost 10%) suggests the potential of the proposed approach.

V. Conclusions

In this work, we proposed a multi-channel approach for audio-based PEP estimation.

The attained preliminary results suggest that signal redundancy leads to better results, in comparison to signal-channel fixed-site methodologies, as the one we proposed in the past.

Although our approach was only evaluated on a two-channel setup, it can be seamlessly applied to a higher number of channels and acquisition sites. We intended to evaluate this possibility in the future.

We will also extend our current dataset, which we believe will allow us to prove the statistical significance of our work. Moreover, we will also evaluate the approach in gender-balanced diseased population.

Finally, we will assess our approach for both PEP and LVET estimation.

References

[1] Oh J. and Tajik J., "The return of cardiac time intervals: the Phoenix is rising", J. of the Am. Col. of Cardiology, vol. 42, no. 8, pp. 1471-1474, 2003.

[2] Gillebert T., van de Veire N., de Buyzere M. and Sutter J., "Time intervals and global function. Use and limita-tions", Eur. Heart J., vol. 25, pp. 2185-2186, 2004.

[3] Weissler A., Harris W. and Schoenfeld C., "Systolic Time Intervals in Heart Failure in Man", Circulation, vol. 37, pp.149-159, 1968.

[4] Werf F., Piessens J. and Kesteloot H., "A comparison of systolic time intervals derived from the central aortic pressure and from the external carotid pulse tracing". Circulation, vol. 51, pp. 310–6, 1975.

[5] Cokkinos D., Heimonas E., Demopoulos J., Haralambakis A., Tsartsalis G. and Gardikas C., "Influence of heart rate increase on uncorrected pre-ejection period/left ventricular ejection time (PEP/LVET) ratio in normal individuals", British Heart Journal, vol. 38, pp. 683-688, 1976.

[6] Lewis R. P., Rittogers S. E., Froester W. F. and Boudoulas H, "A critical review of the systolic time intervals", Circulation, Vol. 56, pp. 146-158, 1977.

[7] Finckelstein S. and Cohn J., "Method and apparatus for measuring cardiac output", US patent 5241966, 1993.

[8] Muehlsteff J., Aubert X. and Schuett M., "Cuffless estimation of systolic blood pressure for short effort bicycle tests: the prominent role of the pre-ejection period", Int. Conf. of the IEEE Engineering in Medicine and Biology Society, 2006.

[9] Couceiro R., Carvalho P., Paiva R.P., Henriques J., Antunes M., Quintal I. and Muehlsteff J., "Beat-to-beat Cardiac Output inference using heart sounds", Proceedings of the Int. Conf. of the IEEE Engineering in Medicine and Biology Society, 2011.

[10] Chan G. S., Middleton P. M., Celler B. G., Wang L. and Lovell N. H., "Automatic detection of left ventricular ejection time from a finger Photoplethysmography pulse oximetry waveform: comparison with Doppler aortic measurement", Physiological Measurement, Vol. 28, pp. 439-452, 2007.

[11] Rietzschel E., Boeykens E., Buyzere M., Duprez D. and Clement D., "A comparison between systolic and dias-tolic pulse contour analysis in the evaluation of arterial stiffness", Hypertension, J. of the American Heart Association, 37:e15, 2001.

[12] Paiva R. P., Carvalho P., Couceiro R., Henriques J., Antunes M., Quintal I. and Muehlsteff J., "Beat-to-Beat Systolic Time-Interval Measurement from Heart Sounds and ECG", Physiological Measurement, Vol. 33, pp. 177-194, 2012.

[13] Sherwood A., Allen M. T. and Fahrenberg J., "Methodological guidelines for impedance cardiography", Psychophysiology, 27:1-23, 1990.

[14] Carvalho P., Paiva R. P., Couceiro R., Henriques J., M. Antunes, Quintal I. and Muehlsteff J., "Comparison of Systolic Time Interval Measurement Modalities for Portable Devices", Proceedings of the Int. Conf. of the IEEE Engineering in Medicine and Biology Society, 2010.

[15] Jiang D., Lu L., Zhang H.-J., Tao J.-H., Cai L.-H., "Music type classification by spectral contrast feature",. Proceedings. 2002 IEEE International Conference on Multimedia and Expo, vol.1, pp. 113-116, 2002.

[16] Henriques J., Carvalho P., Harris M., Antunes M., Couceiro R., Brito M. and Schmidt R., "Assessment of Arrhythmias for Heart Failure Management", phealth2008 - International Workshop on Wearable Micro and Nanosystems for Personalised Health, 2008.

[17] Tavel M., "Clinical Phonocardiography and external pulse recording", Year Book Medical Publishers, Inc., 1967.

Preprint submitted to 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Received April 10, 2015.