Denoising and Segmentation of the Second Heart Sound Using Matching Pursuit

F. Hedayioglu[1], M. G. Jafari[2], S.S. Mattos[3], M. D. Plumbley[2] and M. T. Coimbra[1]

Abstract— We propose a denoising and segmentation technique for the second heart sound (S2). To denoise, Matching Pursuit (MP) was applied using a set of non-linear chirp signals as atoms. We show that the proposed method can be used to segment the phonocardiogram of the second heart sound into its two clinically meaningful components: the aortic (A2) and pulmonary (P2) components.

I. INTRODUCTION

Auscultation is one of the simplest, quickest and most cost effective techniques to listen to heart sounds. It is used to identify and diagnose a large number of heart conditions [1], making it an invaluable cardiac screening tool.

However, cardiac auscultation is a difficult skill to master. Its interpretation is subjective and the heart sounds are difficult to interpret, requiring extensive training and years of experience [2].

Furthermore, heart sounds are highly complex in themselves. To give an idea, there are two main heart sounds S1 (first heart sound) and S2 (second heart sound). It has been recognised that S1 is comprised of up to four components produced during ventricular contraction. S2 is comprised of two main components [3]; the aortic component (A2), which is the sound produced by the closure of the aortic valve, and the pulmonary component (P2) (Figure 1) ,which is the sound produced by the closure of the pulmonary valve.

In addition, some cardiopathies and cardiac conditions, such as pulmonary hypertension can change the amplitude and delay of A2 and P2. Other cardiopathies can change their order of appearance [3], [4]. Therefore the correct identification of these components can be an important indicator for screening and diagnosing some heart conditions and diseases.

Some work has been done to segment the A2 and P2 components of the second heart sound, among them Xu et al [5], [6] model the A2 and P2 components as non-linear narrow-band chirp signals, with a fast decreasing

[1]F. Hedayioglu and M. T. Coimbra, are with the Department of Computer Science, Universidade do Porto, Instituto de Telecomunicações. Email: fheday, mcoimbra @ dcc.fc.up.pt

[2]S. S. Mattos is with Unit of Fetal and Pediatric Cardiology, Royal Portuguese Hospital, Recife, Brazil. E-mail: ssmattos@cardiol.br

[3]M. G. Jafari and M. D. Plumbley, are with the Queen Mary University of London, Centre for Digital Music, London, UK. E-mail: maria.jafari, mark.plumbley@eecs.qmul.ac.uk.
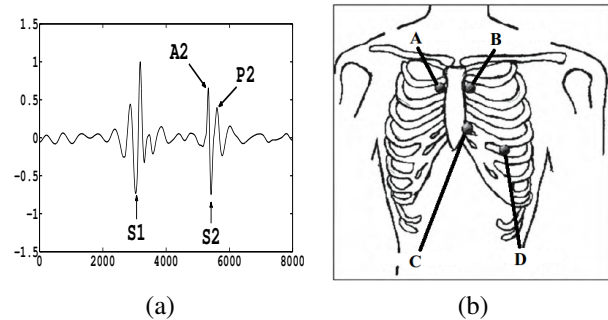
Fig. 1. a) Standard auscultation sites and b) the main components of the heart sounds

instantaneous frequency over time. This method estimate the component with highest energy by using a visually designed 2D mask of the Wigner-Ville distribution to estimate its instantaneous frequency. Then reconstruct the component and subtract it from the signal, repeating this procedure until all components are estimated.

Nigam et al [7] proposed a method for extracting A2 and P2 components by assuming them as statistically independent. They used a non-conventional 4 sensors stethoscope in order to record four simultaneous auscultations and to analyse them using blind source separation. The main advantage of this method is the lower dependence on the A2-P2 time interval.

In [8] the authors also assumed statistical independence of A2 and P2 to perform blind source separation. They took advantage of the fact that in a normal auscultation procedure, the clinician sequentially listens to the heart sounds in four different locations. The periodicity of the heart sounds allowed the clinician to align the recordings of each location and simulate parallel recordings to be able to apply blind source separation and extract the A2-P2 interval.

In this paper we propose a technique to denoise the second heart sound (S2) by using matching pursuit (MP) using physiologically based time-frequency atoms. The second objective of this paper is to propose a method can automatically detect A2 and P2 on the approximated S2 generated by the MP algorithm.

II. MATCHING PURSUIT

Matching Pursuit (MP) is an algorithm used to decompose a signal into a linear expansion of waveforms that are selected from a dictionary of time-frequency functions (called atoms) [9]. MP is a greedy method that finds a representation that is sparse in the dictionary D, i.e. only

a few atoms participate in the approximations. The fact that MP finds a sparse representation of the signal means that it can represent the underlying structures of the signal in a compact representation. Sparse representations are also effective when used for denoise [10].

The MP algorithm represents the signal s as a sum of weighted atoms $\varphi_{\gamma k}$ from a dictionary D, i.e.:

$$s^{(m)} = \sum_{k=0}^{m} \alpha_k \varphi_{\gamma k} \quad (1)$$

Where α_k are the weights of each $\varphi_{\gamma k}$ atom. The algorithm starts from an initial approximation $s^{(0)} = 0$, and a residual $r^{(0)} = s$, it looks for the atom in the dictionary that has the highest dot product with the current residual. Once this atom is found, a scalar multiple of that atom is added so that $s^{(k)} = s^{(k-1)} + \alpha_k \varphi_{\gamma k}$, where $\alpha_k = <r^{(k-1)}, \varphi_{\gamma k}>$ and $r^{(k)} = s - s^{(k)}$ as described in the Algorithm 1.

Algorithm 1 The MP algorithm. The two stop conditions are: the maximum number of iterations $itertLim$, and the minimum error threshold $errorThreshold$.

1. Initialize:
   $s^0 = 0$
   $r^0 = s$
2. repeat:
3. Find $\varphi_{\gamma k} \in D$ that $max <r^{(k-1)}, \varphi_{\gamma k}>$
4. $\alpha_k = <r^{(k-1)}, \varphi_{\gamma k}>$
5. $s^{(k)} = s^{(k-1)} + \alpha_k \varphi_{\gamma k}$
6. $r^{(k)} = s - s^{(k)}$
7. until $(k > iteratLim) \vee (r^{(k)} \leq errorThreshold)$

The approximated signal $s^{(k)}$ is decomposed into a series of time-frequency atoms in decreasing energy order.

$$s^{(m)} = \sum_{k=0}^{m-1} <r^{(k)}, \varphi_{\gamma k}> \varphi_{\gamma k} + r^{(m)} \quad (2)$$

where $r^{(m)}$ is the residual vector after s is approximated by m atoms, and is the projection (weight) of $r^{(k)}$ onto an atom $\varphi_{\gamma k}$.

III. PROPOSED METHOD

A. Pre-processing

In order to discard the frequencies outside the frequency range of the heart sounds, the recordings used in this study were pre-filtered using wavelet threshold denoise using Daubechies 5 level 3 [11], discarding the higher frequency band (above 130Hz). This frequency range corresponds to the low and medium frequency range used to observe the components of the second heart sound in the phonocardiography [3].

B. Signal approximation

After pre-filtering the heart sounds we apply the matching pursuit algorithm.

Since we are looking to match the data with its underlying physiological components. We used, as dictionary D, a non-linear chirp signal model based on the modelling of Xu et al [5], [6] used to model A2 and P2. We assume that these atoms give a good approximation of the signal, since we are making assumptions about the underlying signals that compose S2. The atoms of the dictionary for the MP were made unit vectors, and their equations are as follows:

$$g_{\gamma n}(s, u, f_1, f_2, t) = \frac{1}{\sqrt{(s)}} a(t) \sin(\frac{f_1 t}{10} + 2f_2 \sqrt{\frac{t}{10}} - 2f_2) \quad (3)$$

$$a(t) = (1 - e^{-\frac{t}{8}})e^{-\frac{t}{16}} \sin(\frac{\pi t}{60}) \qquad 0 \leq t \leq 60ms \quad (4)$$

Where s is the scale, u is the displacement (or shift), f_1 is the atom's lowest frequency and f_2 is the atom's highest frequency.

For each S2, we performed MP using two atoms as stop-condition, therefore, creating a signal that is an approximation of S2 (Figure 2).
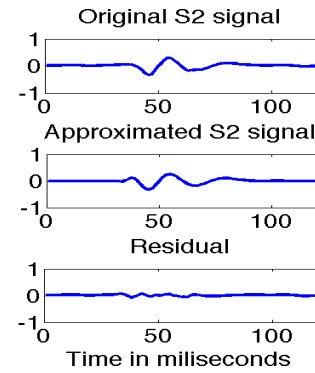

Fig. 2. Recorded S2 (top), approximated S2 by MP (middle) and its residual (bottom)

C. Automatic detection of A2 and P2

The detection of A2 and P2 happens in the approximated signal $s^{(k)}$. Here we are trying to reproduce what the clinician is looking for: the two highest and more distinguished peaks in the second heart sound. Since we are assuming a healthy person, we can assume that A2 comes before P2. For each S2, we execute the Algorithm 2.

IV. EXPERIMENTAL RESULTS

A. Recording and pre-processing

The experiment had ethical approval from Queen Mary University of London, and the recruited subject was a 31 years old student, with no cardiopathy. The recordings were made using the Thinklabs ds32a+ digital stethoscope recording at 44.1KHz and with 16 bits of resolution, with all filters turned off.

To reduce computational complexity, the recordings were downsampled to 1050Hz. The filtering was performed as described in section III-A.

Algorithm 2 pseudo-code for detecting A2 and P2.

1. $dif = \frac{\partial S^{(k)}(t)}{\partial t}$
2. Look for changes in the signal of dif(peaks)
3. For all positive peaks:

   $p(i) = \int_a^b s^{(k)}(t)\, dt$, where $\begin{cases} S^{(k)}(a) = 0 \wedge a < i \\ S^{(k)}(a) = 0 \wedge b > i \end{cases}$
4. tmp = sort(p)
5. if tmp(0) is earlier than tmp(1)

   A2 = tmp(0)

   P2 = tmp(1)
6. $else$

   A2 = tmp(1)

   P2 = tmp(0)
7. endIf
8. return A2 and P2

B. Manual annotation

In a normal procedure, the clinicians listen to the heart sounds and try to identify its components: the first and second heart sound, the A2 and P2 components, the presence of murmurs, etc. This is usually done by ear only, but can be better measured by using a phonocardiogram (the graphical representation of the heart sound). In this case, a clinician can visually identify and measure A2 and P2. By doing so, he can then observe some important features such as the amplitude of A2 and P2, the amplitude of A2 in relation to the amplitude of P2, the time interval between them, etc.

In normal patients, A2 is defined as being the earliest highest positive peak in S2, and P2 is the most prominent peak that follows A2. Since the duration of these components is quite short (Table I), it is difficult to manually annotate them.

We chose to make the annotations by identifying all the intervals in the recordings where A2 and P2 could be found, as shown in Figure 3. Then we compute the precise parameters of A2 and P2 (amplitude and position in time) by finding the positive peaks in the recording inside the identified intervals and measure its amplitude and time. Based on these amplitudes, we calculate the mean amplitude, the lowest observed amplitude and highest observed amplitude of each observed component. Later on, these annotations will be used to compare the amplitudes and position in time of the annotated A2 and P2 with the automatically detected ones. If a heart sound component is detected outside the annotated area, then that heart sound is discarded. On each auscultation site (Figure 1), the number of annotated S2 was 128. In total, 512 S2 were annotated. These annotations were made to be as short as possible. On Table I we can see that these annotations are quite short, since the average duration of the S2 sound is about 0.06 seconds [5].

V. EXPERIMENTAL RESULTS

It was considered for the analysis of the results only the A2 and P2 that were detected inside the annotated interval on each of the 512 S2 recorded. By analysing Table II, we
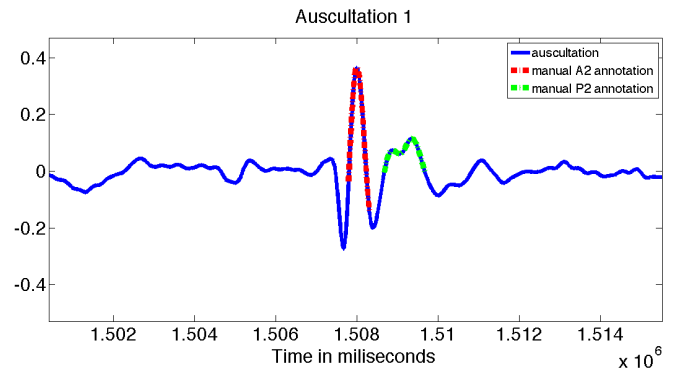

Fig. 3. S2 and the manually annotated regions where A2 (red) and P2 (green) can be found

TABLE I

SIZE IN SECONDS OF THE ANNOTATIONS FOR A2 AND P2

	A2	P2
Mean	0.0131	0.0274
Min	0.0086	0.0114
Max	0.0200	0.0410

can see that the auscultation site D had the most number of discarded components. This high percentage of discarded components in the auscultation site D demonstrate that at this site, the recording is poor, if compared with the other sites. A possible explanation resides in the fact that this auscultation site is the most distant from the pulmonary valve (that is closer to the auscultation sites A and B).

By analysing Table III and Table IV we can see that the mean amplitude of both components through the recordings are quite close. Their standard deviation is also similar: in both cases, the amplitude of P2 varies more than A2 - This is expected in normal subjects, since the amplitudes of A2 and P2 are related to the systemic and pulmonary pressure respectively and the pulmonary pressure changes more than the systemic due to the breathing cycle [4]. Also, normal subjects have $A2 > P2$, this is also confirmed on both tables [4].

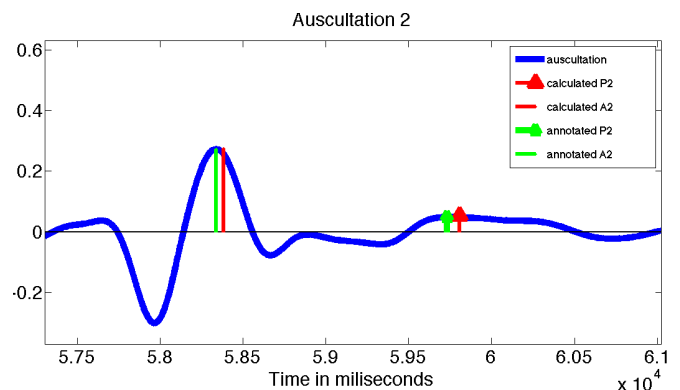The high correlation between the approximated S2 and the

Fig. 4. Detected and annotated A2 and P2 on the recorded heart sound

TABLE II

THE NUMBER OF REJECTED COMPONENTS. IN PARENTHESIS IS THE
PERCENTAGE RELATIVE TO THE TOTAL OF ANNOTATED COMPONENTS.

	Site A	Site B	Site C	Site D
Discarted A2	0(0%)	0(0%)	0(0%)	11(8.6%)
Discarted P2	25 (19.5%)	21(16.4%)	21(16.4%)	60(46.9%)

TABLE III

THE MEAN AND STANDARD DEVIATION OF THE AMPLITUDES OF THE
ANNOTATED A2 AND P2 ACROSS THE 128 S2 IN THE RECORDED
SIGNAL.

	Site A	Site B	Site C	Site D
mean A2	0.3324	0.3046	0.3722	0.1497
σ	0.0437	0.0477	0.0594	0.0393
mean P2	0.0801	0.0813	0.0770	0.0651
σ	0.0245	0.0189	0.0202	0.0230

TABLE IV

THE MEAN AND STANDARD DEVIATION OF THE AMPLITUDES OF THE
AUTOMATICALLY DETECTED A2 AND P2 IN THE SIGNAL APPROXIMATED
BY MP.

	Site A	Site B	Site C	Site D
mean A2	0.2787	0.2904	0.3274	0.1278
σ	0.0539	0.0508	0.0528	0.0372
mean P2	0.0858	0.0916	0.0898	0.0562
σ	0.0244	0.0232	0.0240	0.0249

TABLE V

THE APPROXIMATED S2 CORRELATION WITH THE ORIGINAL S2 IN
PERCENTAGE, BY AUSCULTATION SPOT.

	Site A	Site B	Site C	Site D
Mean	92.55%	95.00%	93.70%	88.71%
Min	81.61%	80.09%	83.62%	74.64%
Max	97.19%	98.33%	97.82%	96.66%

TABLE VI

THE APPROXIMATED S2 MEAN SQUARED ERROR BY AUSCULTATION
SPOT.

	Site A	Site B	Site C	Site D
A2	0.0315	0.0079	0.0187	0.0428
P2	0.2181	0.0886	0.3038	0.1132

original ones shows that the approximated signal is a good approximation of the real one (Table V)

The average squared error between the detected A2, P2 and the annotated A2, P2, per auscultation site (Table VI) shows that better estimations of P2 are made at the site B (second left intercostal space): this is, indeed, the best site to identify the pulmonary component [4].

VI. DISCUSSION

We validated the results by comparing them with data previously annotated. From all sites for auscultation, the site D was the one that showed poorer results. We assumed that it is due to the fact that it is too far from the aortic and pulmonary arteries. This is supported by the fact that this site had the lowest correlation with the original recording among all sites (Table V).

For future work, we plan to work on optimizations for processing time and lower the memory usage by using information of the time-frequency analysis of the recorded heart sound to lower the search space of the MP.

Despite being able to reasonably represent S2, our atoms seems to underestimate the value of A2 and overestimate P2. This may be due to the amplitude function in equation (4) not decaying quickly enough to reflect these changes.

Another future work will be the integration of this technique into a digital stethoscope to be used in a real clinical environment

VII. CONCLUSION

A new technique to denoise the second heart sound by using MP with non-linear chirp signals was proposed. We also demonstrated that the proposed method can successfully automatically segment A2 and P2.

This new technique has the potential to be used in a real clinical environment as a tool for improving the measurement of physiologically relevant components of heart sounds.

REFERENCES

[1] S. Yuenyong and W. Kongpravechnon and K. Tungpimolrut and A. Nishihara, "Automatic heart sound analysis for tele-cardiac auscultation", Proc. ICROS-SICE Int. Joint Conference, pp. 1599-1604, 2009.

[2] U. Alam, O. Asghar, S. Khan, S. Hayat and R. Malik, "Cardiac auscultation: an essential clinical skill in decline", British Journal of Cardiology, pp. 8-10, Vol.17, 2010.

[3] A. Luisada, "From auscultation to phonocardiography",Mosby, 1965.

[4] M. E. Tavel, "Clinical phonocardiography and external pulse recording",Year book medical publishers, 1985.

[5] J. Xu, L. Durand, and P. Pibarot, "Nonlinear transient chirp signal modeling of the aortic and pulmonary components of the second heart sound", IEEE Transactions on Biomedical Engineering, pp. 1328-1335, Vol.47, 2000.

[6] J. Xu, L. Durand, and P. Pibarot, "Extraction of the aortic and pulmonary components of the second heart sound using a nonlinear transient chirp signal model", Transactions on biomedical engineering, pp. 277-283, Vol.48, 2001.

[7] V. Nigam and R. Priemer, "A procedure to extract the aortic and the pulmonary sounds from the phonocardiogram", In Proceedings of the 28th IEEE EMBS Annual International Conference, pp. 5715-5718, 2006.

[8] F. Hedayioglu, M. Jafari, S. Mattos, M. Plumbley, and M. Coimbra, "Separating sources from sequentially acquired mixtures of heart signals", Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, pp. 653-656, 2011.

[9] S.G. Mallat, and Z. Zhang, "Matching pursuits with time-frequency dictionaries", IEEE Transactions on Signal Processing, pp. 3397-3415, Vol.41, 1993.

[10] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries", Image Processing, IEEE Transactions on, pp. 3736-3745, Vol. 15, 2006.

[11] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis", IEEE Transactions on Information Theory, pp. 961-1005, Vol.36, 1990.