Source Separation of the Second Heart Sound via Alternating Optimization

Francesco Renna[1], Mark D Plumbley[2], Miguel Coimbra[3]

[1] Instituto de Telecomunicações, Faculdade de Ciências da Universidade do Porto, Portugal
[2] University of Surrey, UK
[3] INESC TEC, Faculdade de Ciências da Universidade do Porto, Portugal

Abstract

A novel algorithm to separate S2 heart sounds into their aortic and pulmonary components is proposed. This approach is based on the assumption that, in different heartbeats of a given recording, aortic and pulmonary components maintain the same waveform but with different relative delays, which are induced by the variation of the thoracic pressure at different respiration phases. The proposed algorithm then retrieves the aortic and pulmonary components as the solution of an optimization problem which is approximated via alternating optimization.

The proposed approach is shown to provide reconstructions of aortic and pulmonary components with normalized root mean-squared error consistently below 10% in various operational regimes.

1. Introduction

Measuring directly the pulmonary artery pressure (PAP) requires invasive procedures, as cardiac catheterization, which has recently motivated research efforts aiming to infer information regarding the PAP from non-invasive and less costly measurements. One solution in this direction is represented by the use of Doppler echocardiography to estimate PAP, which is already included in clinical use. On the other hand, this technique presents a series of limitations, as the need of a specialized doctor to perform the exam and high error levels for patients with chronic obstructive pulmonary disease and patients with normal to elevated PAP [1, 2].

A convenient alternative to Doppler echocardiography in inferring information about the PAP from non-invasive measurements is represented by cardiac auscultation, due to the low cost and simplicity of performing such exam. In particular, the analysis of the second heart sound (S2) can be used to estimate the PAP, by noting that the S2 is originated by the closure of the aortic and pulmonary valves, and their vibrations. In particular, the sound produced by the closure of the aortic valve is called the A2 component, whereas the sound generated by the closure of the pulmonary valve is called the P2 component. Accurate estimation of the PAP requires to handle the A2 and P2 components of S2 sounds separately [3]. On the other hand, the separation of the A2 and P2 components of S2 sounds represents a very challenging problem, due to their large overlap in the time and frequency domain and their morphological similarity. Moreover, the A2 and P2 components cannot be modeled as independent signals, due to the interaction of the mechanical processes that originate them. For these reasons, standard methods for blind source separation such as independent component analysis (ICA) [4] and morphological component analysis (MCA) [5] are not suitable to solve such separation problem.

Motivated by these observations, several separation approaches specifically tailored to the characteristics of S2 signals have appeared in the literature in recent years. Some of these are based on the assumption that A2 and P2 components can be modeled via predetermined *ad hoc* waveforms, whose parameters are fitted to the observed S2 sounds. The selected models include transient nonlinear chirp models [1], Gaussian chirplets [6], and Gaussian windowed sinusoids [7].

Other approaches have recently emerged that refrain from adopting a predetermined waveform model, using instead the different dynamical behavior of A2 and P2 components in different S2 sounds to perform separation. In particular, the work by Tang *et al.* [8] leverages the fact that A2 components usually occupy the same position in different S2 sounds, whereas P2 components experience a shift according to the respiration phase of each recorded S2 sound (i.e., P2 components experience a more significant delay from A2 components during inspiration phases than during expiration phases [9]). Therefore, assuming that A2 and P2 components keep approximately the same form (apart from time shifts) in successive S2 sounds, the A2 components are estimated by averaging over the collected S2 sounds and the P2 components are obtained by subtraction. Similar assumptions on the dynamic behavior of A2 and P2 components have been adopted by the approach presented in [10], which attempts to recover A2 and P2 components using a joint multivariate Gaussian mixture

model (GMM) spanning samples extracted from S2 sounds at different heartbeats. However, this approach requires the availability of a training set with already isolated A2 and P2 components, which is difficult to collect in common clinical practice.

In this work, starting from similar assumptions on the dynamic behavior of A2 and P2 components in different heartbeats, we propose a novel separation algorithm specifically tailored for the separation of the A2 and P2 components of S2 heart sounds. The proposed method allows for unsupervised inference of the specific shape of the waveforms associated to the A2 and P2 components from the data, without defining predetermined, *ad hoc* waveforms. The A2 and P2 components are obtained by defining the separation problem as an optimization problem on the shapes and positions of the corresponding waveforms, which is solved using an alternating optimization method.

In particular, this work proposes a novel separation approach, specifically tailored for S2 sounds separation, that involves the following contributions:

1. The proposal of a novel, non-parametric separation algorithm for the S2 sound based on alternating optimization which leverages the physiological analysis of how A2 and P2 components are generated in different heartbeats;

2. The analysis of the separation performance of the proposed algorithm and comparison with other separation approaches based on similar physiological assumptions on a set of synthetically generated heart sounds.

2. Proposed method

In this section, the proposed approach for the separation of the aortic and pulmonary components of S2 sounds is described in details.

2.1. Signal model and setup

We are presented with the observation of a series of M S2 sounds obtained from a recording of a given patient. Each S2 sound in the recording consists of the superposition of two components, the A2 and the P2 component. The observed signals are assumed to be sampled with sampling period T_s. Moreover, N samples are collected for each S2 sound.

We assume that the A2 and P2 components in different S2 sounds of the same recording have the same shape, but they have different delays, so that we can model the observed S2 sounds as follows:

$$s_m(tT_s) = a(tT_s - \tau_{A,m}) + p(tT_s - \tau_{P,m}), \quad (1)$$

for $m = 1, \dots, M$ and $t = 0, \dots, N-1$.

Then, the objective of the considered source separation method is to recover $a(tT_s - \tau_{A,m})$ and $p(tT_s - \tau_{P,m})$ from the observation of $s_m(tT_s)$, for $m = 1, \dots, M$.

The proposed separation method is based on the formulation of a least-squares optimization problem whose solution is approximated via an alternating iterative method. The considered optimization problem is the following:

$$[\hat{a}(\cdot), \hat{p}(\cdot), \{\hat{\tau}_{A,m}\}, \{\hat{\tau}_{P,m}\}]$$
$$= \arg \min_{a(\cdot), p(\cdot), \{\tau_{A,m}\}, \{\tau_{P,m}\}} \sum_{m=0}^{M-1} \sum_{t=0}^{N-1} |s_m(tT_s)$$
$$- a(tT_s - \tau_{A,m}) - p(tT_s - \tau_{P,m})|^2, \quad (2)$$

and it aims to minimize the mean-squared error (MSE) between the S2 observations and the superposition of the separated A2 and P2 components.

After an initialization step, the optimal solution (2) will be approximated by an alternating iterative algorithm which alternates between finding the values $\{\tau_{A,m}\}, \{\tau_{P,m}\}$ that minimize the MSE, for a given pair of waveforms $a(\cdot), p(\cdot)$ and finding the signals $a(\cdot), p(\cdot)$ which minimize the MSE for given sets of delays $\{\tau_{A,m}\}, \{\tau_{P,m}\}$, for a fixed number I of iterations.

2.2. Initialization

The initial estimates of waveforms and delays associated to A2 and P2 components are obtained by applying the method described in [8], which computes an estimate of the A2 component by simply averaging the observed S2 sounds. Therefore, we can write the initial estimate $a^{(0)}(tT_s)$ as:

$$a^{(0)}(tT_s) = \frac{1}{M} \sum_m s_m(tT_s), \quad (3)$$

where we have introduced the use of the superscript $(\cdot)^{(i)}$ to denote the estimate of a given delay/function at the i-th iteration of the algorithm. According to the model used in [8], the delays associated to the A2 components are the same in all S2 sounds. Therefore, without loss of generality, we can assume in our initialization $\tau_{A,m}^{(0)} = 0$ for all $m = 1, \dots, M$.

Then, the values of $p^{(0)}(\cdot)$ and $\tau_{P,m}^{(0)}$ are obtained by considering the signals

$$s_m(tT_s) - a^{(0)}(tT_s).$$

Such signals are then aligned via a maximum correlation method and averaged in order to obtain $p^{(0)}(tT_s - \tau_{P,0}^{(0)})$. Finally, the remaining delays $\tau_{P,m}^{(0)}$ are determined computing the correlation between $p^{(0)}(tT_s - \tau_{P,0}^{(0)})$ and the signals $s_m(tT_s) - a^{(0)}(tT_s)$.

2.3. Delay estimation

Given the estimates $a^{(i)}(\cdot), p^{(i)}(\cdot)$ of the A2 and P2 waveforms obtained from the i-th iteration of the algorithm, the corresponding delay estimates $\{\tau_{A,m}^{(i)}, \tau_{P,m}^{(i)}\}$ are obtained as the solution of the following problem:

$$
\begin{aligned}
&[\{\tau_{A,m}^{(i)}\}, \{\tau_{P,m}^{(i)}\}] \\
&= \arg \min_{\{\tau_{A,m}\}, \{\tau_{P,m}\}} \sum_m \sum_t \Big| s_m(tT_s) \\
&\quad - a^{(i)}(tT_s - \tau_{A,m}) - p^{(i)}(tT_s - \tau_{P,m}) \Big|^2, \quad (4)
\end{aligned}
$$

where it is straightforward to note that the optimization can be performed independently for the different indexes m.

The solution of (4) is obtained using a greedy approach that computes the values of the MSE for all possible pairs of values $\{\tau_{A,m}^{(i)}, \tau_{P,m}^{(i)}\}$ that are multiple of the sampling rate T_s and that reside in a physiologically feasible range. In particular, on noting that the time split between A2 and P2 components usually ranges between 10 ms and 50 ms [8], the optimal delay search considers delay values in the interval $[-100, 100]$ ms.

2.4. Waveform estimation

Given the current delay estimates $\tau_{A,m}^{(i)}, \tau_{P,m}^{(i)}$, and the observations $s_m(tT_s)$, the signal waveforms associated to the $(i+1)$-th iteration, $a^{(i+1)}(\cdot), p^{(i+1)}(\cdot)$ are given by the solution the following optimization problem:

$$
\begin{aligned}
&[a^{(i+1)}(\cdot), p^{(i+1)}(\cdot)] \\
&= \arg \min_{a(\cdot), p(\cdot)} \sum_m \sum_t \Big| s_m(tT_s) \\
&\quad - a(tT_s - \tau_{A,m}^{(i)}) - p(tT_s - \tau_{P,m}^{(i)}) \Big|^2. \quad (5)
\end{aligned}
$$

On assuming that the admissible delay estimates are always obtained as multiples of the sampling period T_s, the problem in (5) can be cast as a standard linear least-squares problem, which is overdetermined when $M > 2$. The problem can be expressed in matrix form as follows:

$$
\begin{bmatrix} \mathbf{a}^{(i+1)} \\ \mathbf{p}^{(i+1)} \end{bmatrix} = \arg \min_{\mathbf{a}, \mathbf{p} \in \mathbb{R}^N} \left\| \mathbf{s} - \mathbf{X}^{(i)} \begin{bmatrix} \mathbf{a} \\ \mathbf{p} \end{bmatrix} \right\|_2^2, \quad (6)
$$

where the vector $\mathbf{a} \in \mathbb{R}^N$ contains the samples of the signal $a(\cdot)$, the vector $\mathbf{p} \in \mathbb{R}^N$ contains the samples of the signal $p(\cdot)$, and the vector $\mathbf{s} \in \mathbb{R}^{M \cdot N}$ contains the samples of the signals $s_m(\cdot)$ for $m = 0, \ldots, M - 1$. The matrix $\mathbf{X}^{(i)} \in \mathbb{R}^{M \cdot N \times 2N}$ is formed by $M \times 2$ blocks of dimension $N \times N$, each containing a "shifted version" of the $N \times N$ identity matrix, where the values of such shifts depend on the delays $\tau_{A,m}^{(i)}$ and $\tau_{P,m}^{(i)}$.
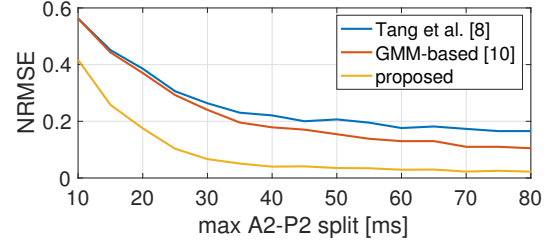

Figure 1. Separation performance over synthetic S2 sounds. Separation NRMSE vs. maximum A2-P2 split.

When the matrix $\mathbf{X}^{(i)}$ has full column rank, the corresponding linear least-squares problem (6) has a unique solution which is given by

$$
\begin{bmatrix} \mathbf{a}^{(i+1)} \\ \mathbf{p}^{(i+1)} \end{bmatrix} = (\mathbf{X}^{(i)})^\dagger \mathbf{s} = \left((\mathbf{X}^{(i)})^{\mathrm{T}} \mathbf{X}^{(i)} \right)^{-1} (\mathbf{X}^{(i)})^{\mathrm{T}} \mathbf{s}, \quad (7)
$$

where $(\cdot)^\dagger$ is the Moore-Penrose pseudoinverse.

3. Numerical results

In this section, numerical results obtained with synthetic heart sound signals are shown in order to validate the proposed method and to compare its separation performance with that of the methods described in [8] and [10], which rely on a similar set of assumptions regarding the variability of A2 and P2 components in different S2 sounds.

These results are obtained considering synthetic signals generated using the model described in [7]. In particular, S2 sounds are obtained as the superposition of A2 and P2 components modeled via Gaussian windowed sinusoids. In all the reported results, the proposed alternating optimization method is run with a fixed number of iterations, $I = 10$, and separation results are averaged over 40 randomly generated heart sound recordings, each containing 40 heartbeats.

First, the separation algorithms are tested when varying the maximum value of the A2-P2 split, i.e., the distance between the two components in a given S2 sounds. In particular, Fig. 1 contains the normalized root mean-squared error (NRMSE) values obtained by the proposed separation algorithm and the methods in [8] and [10] when generating synthetic S2 sounds for each heart sound recording with A2-P2 splits drawn from a uniform distribution over the interval $[0, \Delta_{\max}]$ ms, for different values of $\Delta_{\max}$.

As expected, all the considered methods achieve better separation results when the temporal separation between the components of S2 sounds is larger. On the other hand, the proposed approach clearly provides significantly more accurate reconstructions for all considered values of $\Delta_{\max}$.

The robustness of the considered separation approaches is also tested for the case when the recorded S2 sounds do
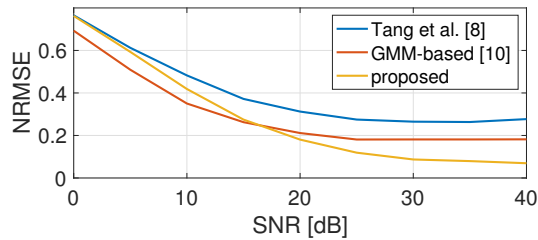
Figure 2. Separation performance over synthetic S2 sounds. Separation NRMSE vs. SNR.

not perfectly satisfy the modeling assumptions described in Section 2.1. In particular, the generated A2 and P2 components are contaminated with independent additive white Gaussian noise (AWGN) at different signal-to-noise ratio (SNR) levels. Note that the presence of such distortion implies that A2 and P2 components in different heartbeats of the same recording are not exactly modeled via shifted versions of the same waveforms. Fig. 2, reports the separation NRMSE obtained with the two considered separation algorithms for different SNR values and for $\Delta_{max} = 40$ ms.

Also in this case, the proposed approach guarantees better reconstruction performance than the method in [8] for all the range of considered SNR values. This results is possibly explained by the increased flexibility of the proposed approach modeling assumptions which do not impose perfect alignment of A2 sounds in all the observed S2 recordings. On the other hand, for low SNR values, the approach in [10] achieves better performance, but this is obtained at the expense of the use of annotated training data to fit the GMM used for reconstruction.

4. Conclusion

In this work, we have presented an unsupervised source separation algorithm specifically designed to identify A2 and P2 components from S2 recordings. The proposed approach has the advantage of refraining from using predefined waveform shapes to fit aortic and pulmonary components. The reconstruction of A2 and P2 components is defined in terms of an optimization problem whose solution is approximated via alternating optimization.

The proposed separation method is shown to provide more accurate and robust reconstructions of aortic and pulmonary components than other approaches formulated on similar premises, when tested over synthetic heart sounds.

Acknowledgements

References

[1] Xu J, Durand LG, Pibarot P. A new, simple, and accurate method for non-invasive estimation of pulmonary arterial pressure. Heart 2002;88(1):76–80.

[2] Smith R, Ventura D. A general model for continuous non-invasive pulmonary artery pressure estimation. Computers in Biology and Medicine 2013;43(7):904–913.

[3] Guedes R, Carvalho HC, Castro A. Heart sound analysis for blood pressure estimation. In Advanced Methodologies and Technologies in Medicine and Healthcare. IGI Global, 2019; 98–109.

[4] Comon P, Jutten C. Handbook of Blind Source Separation: Independent Component Analysis and Applications. 1st edition. Orlando, FL, USA: Academic Press, Inc., 2010.

[5] Bobin J, Starck JL, Fadili J, Moudden Y. Sparsity and morphological diversity in blind source separation. IEEE Transactions on Image Processing Nov 2007;16(11):2662–2674.

[6] Popov B, Sierra G, Durand LG, Xu J, Pibarot P, Agarwal R, Lanzo V. Automated extraction of aortic and pulmonary components of the second heart sound for the estimation of pulmonary artery pressure. In The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, volume 1. IEEE, 2004; 921–924.

[7] Sæderup RG, Hoang P, Winther S, Bøttcher M, Struijk J, Schmidt S, Østergaard J. Estimation of the second heart sound split using windowed sinusoidal models. Biomedical Signal Processing and Control 2018;44:229–236.

[8] Tang H, Chen H, Li T. Discrimination of aortic and pulmonary components from the second heart sound using respiratory modulation and measurement of respiratory split. Applied Sciences 2017;7(7):690.

[9] Amit G, Shukha K, Gavriely N, Intrator N. Respiratory modulation of heart sound morphology. American Journal of Physiology Heart and Circulatory Physiology 2009; 296(3):H796–H805.

[10] Renna F, Coimbra M. Source separation of the second heart sound using Gaussian mixture models. In 2019 Computing in Cardiology (CinC). IEEE, 2019; 1–4.

Address for correspondence:

Name: Francesco Renna
Full postal address: Rua do Campo Alegre 1021/1055,
4169-007 Porto, Portugal
E-mail address: frarenna@dcc.fc.up.pt