

DETECTION OF THE S2 SPLIT USING THE HILBERT AND WAVELET TRANSFORMS

S. Santos, P. Carvalho, R. P. Paiva and J. Henriques

Department of Informatics Engineering
Faculty of Sciences and Technology
University of Coimbra
Pinhal de Marrocos 3030-290 Coimbra
e-mail: sdsantos@student.dei.uc.pt, {carvalho, ruipedro, jh}@dei.uc.pt

Keywords: biosignal processing, S2 split, Hilbert transform, wavelet transform

Abstract. *Pulmonary hypertension (PH) is a serious heart condition that is difficult to diagnose in ambulatory settings. Heart sounds is one of the most relevant diagnosis signal in this context. Usually, PH leads to wide S2 split between the aortic valve close sound (A2) and the pulmonary valve close sound (P2). Even though the progresses made, there isn't still an automatic and non-supervised way to effectively measure the S2 Split. Results based on the wavelet transform, Wigner-Ville distribution and Chirp models are presented in the literature. However, most methods require human intervention and depend on parameters with high intra and inter-patient variability, showing statistically low significant results. Furthermore, most of these methods are evaluated in synthesized signals or non-human heart sounds. In this paper three approaches for S2 Split detection are presented and compared based on the onset of the A2 and P2 components. The approaches are based on the Hilbert transform's instantaneous amplitude (IA) and frequency (IF), and on the continuous wavelet transform IF (CWT). Results from a 226 manually annotated S2 sound database will be presented and discussed. The best approach, a combination of the Hilbert transform IA and IF, obtained an average error of $3.4 \pm 2.8\text{ms}$ (correlation $\rho = 0.99$) on the detection of the A2 onset, and an error of $10.3 \pm 11.2\text{ms}$ (correlation $\rho = 0.99$) on the detection of the P2 start instant.*

1 INTRODUCTION

Heart sounds derive from the cumulative effect of a series of dynamic events inside the thoracic cavity. Heart chambers contraction and relaxation, valve movements and the blood flow are the most significant of those events. Cardiac auscultation is the method for direct listening of heart sounds to screen pathologies, usually performed by doctors using analogue stethoscopes on a patient's chest. Phonocardiography consists on the recording of heart sounds to enable more detailed analysis. In the case of digital recordings, the sounds signals can be computer analyzed, in real-time or off-line, through signal processing algorithms.

The heart sound of a heart beat is composed by a set of small sound segments. The most predominant are S1 and S2, also called normal heart sounds since they are audible in every person. There are two extra heart sounds called S3 and S4 that may be present in healthy children or trained athletes, but can be a result of congestive heart failure or ventricular dilatation in subjects over 35 years old. S1 is the first sound of the heart beat and it is believed to originate mainly from the closure of the mitral and tricuspid valves. This work studies the properties of S2, also known as the second heart sound.

S2 results from two main heart events: the closure of the aortic and pulmonary valves. These produce two high-frequency components, known as A2 and P2 respectively. In healthy persons, the aortic valve closes before the pulmonary valve with a certain time interval. This time difference between A2 and P2 is commonly named the S2 split. Its duration varies with the respiration cycle. During the expiration it is usually less than 15ms; when in the inspiration it widens to 30ms to 80ms [1].

The lack of this variation, also called fixed splitting happens in certain pathologies like atrial septal defect or pulmonary stenosis. A longer S2 split could mean the presence of pulmonary hypertension, the third most common cardiac disorder [2]. The splitting interval can also be used to estimate the Pulmonary Arterial Pressure (PAP), an important metric for the screening of cardiac disorders. Finally, a reverse S2 split or paradoxical splitting, is the term given to the closure of the pulmonary valve prior to the aortic valve. This is a consequence of some heart pathologies like the Left Bundle Branch Block [3].

S2 Split detection has been attempted using several signal and pattern recognition techniques. A nonlinear transient chirp signal model for the A2 and P2 components was proposed in [4]. With them, and after identifying their instantaneous frequency (IF), A2 and P2 can be extracted from a S2 recording. The Wigner-Ville Distribution was used to estimate the IF. However, a manually filter is used to smooth the distribution results. In [5] the blind source separation algorithm Independent Component Analysis (ICA) was tested. This algorithm places a constraint: to separate N source signals, at least N simultaneous recordings of the composition are needed. Therefore, at least two recordings are required for separating the A2 and P2 components from the S2 signal, assuming no noise is present.

Debbal et al. [6] used the discrete wavelet transform to detect the best S2 split of the

sound signal. The Daubechies wavelet was applied and a study was made on which level the split is best detected, according to heart disorders present in the S2 sound sample. The goal of this work was to determine if the patient presented variable duration splitting, a normal condition, or fixed duration splitting, an indicator of heart pathologies. In this paper, three non-supervised approaches for determining the start instant of the A2 and P2 components are presented and compared in sections 2 and 3, respectively. Main conclusions are discussed in the final section.

2 METHODS

2.1 Hilbert transform based method

It was observed that the S2 sound envelope might be a good feature for separating the two S2 components and estimating their amplitude peaks, since A2 and P2 should exhibit well defined IA signatures. We find the signal envelope by retrieving the absolute value of the analytical signal $\tilde{u}(t) = u(t) + j \cdot H(u(t)) = A(t) \cdot e^{j\psi(t)}$, where $u(t)$ is the real valued signal, $H(u(t))$ represents the Hilbert transform computed by equation (1) and $A(t)$ and $\psi(t)$ are the instantaneous amplitude and frequency of the signal. Figure 1 exemplifies the results achieved.

$$H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(t - \tau)}{\tau} d\tau \quad (1)$$

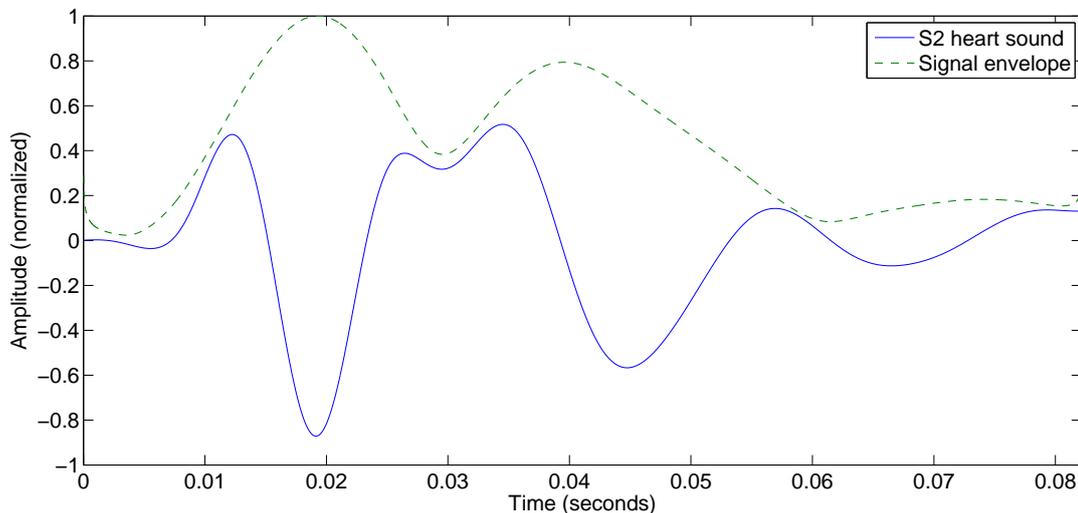


Figure 1: S2 heart sound signal in solid line and its amplitude envelope in dashed line.

To facilitate the component identification, we needed a method for separating the signal envelope in the A2 and P2 segments. Theoretically, the IA of the S2 sound should be a

bimodal signal, since A2 and P2 are two rapidly decaying signals in time with respect to amplitude. The Otsu [7] binarization method, a technique usually applied to distinguish histogram regions, was used to estimate the valley between the two envelopes.

Otsu’s binarization method selects a good threshold point in the signal by maximizing the inter-class variance, a faster and proved method to calculate the intra-class variance minimum. Originally used for thresholding image histograms, it was here applied to estimate the optimum point between the A2 and P2 sound signals. The inter-class variance is calculated with equation (2), where the weights ω_i are the probabilities of the two classes, μ_i the means, and t a iterative threshold that separates the two classes. In this case, ω_i would be the proportion of amplitude area for a given class, and μ_i the weighted mean of each class time instant with its amplitude. The IA signal is normalized before this operation.

$$\sigma_b^2 = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2 \quad (2)$$

The middle point obtained using the above methods is usually located outside an amplitude valley, its optimum place. Therefore, a simple technique was developed to correct its position. The minimum dip in a given region around the previous middle point (the region radius currently used is of 25ms) is selected. Figure 2 presents a distinguishable example of the technique result.

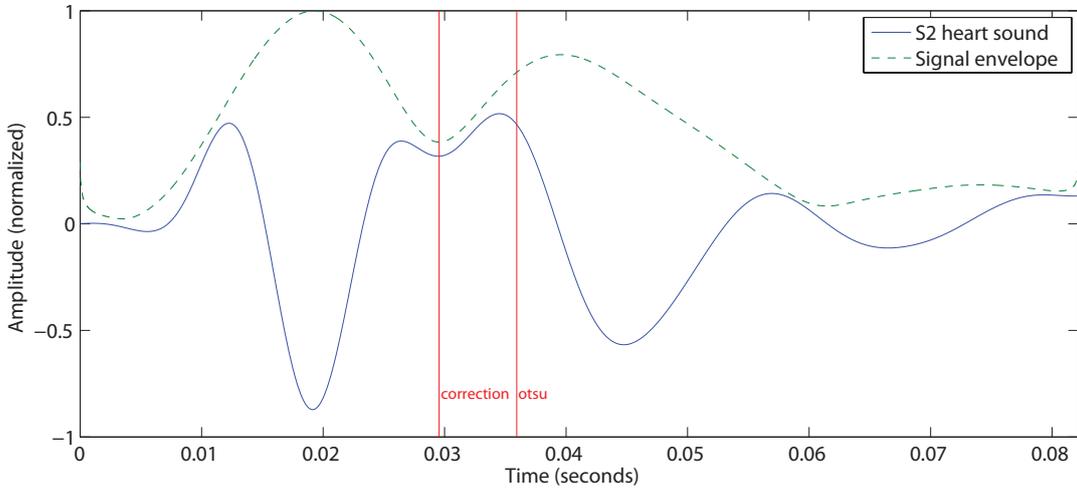


Figure 2: Amplitude valley Otsu’s estimation and its correction. The solid line represents a S2 sounds signal, and the dashed line its envelope.

We identified the amplitude peaks for the A2 and P2 components similarly to the middle point. First, the rightmost peak on the region left of the middle point, and the leftmost peak on the region right of the middle point, were identified as A2 and P2 amplitude peaks, respectively. Later, given the assumption that the S2 sound region was already

correctly segmented, the peaks were selected as the maximum values of each region. This change improved results when the corrected middle point still was not placed on the valley before the P2 sound occurrence. A simple example of amplitude peak detection is shown in Figure 3.

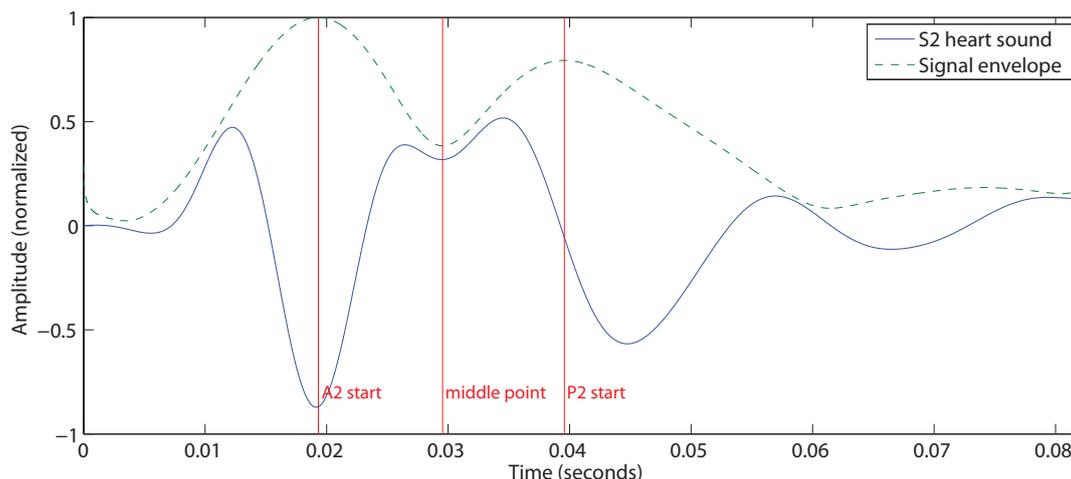


Figure 3: A2 and P2 amplitude peak identification. The solid line represents a S2 sounds signal, and the dashed line its envelope.

To estimate the start of the A2 sound, we approximate the left slope of the A2 amplitude peak to a line with negative derivative. The line is calculated using two focal points with amplitude at 40% and 60% of the peak's amplitude. The time instance where the line reaches zero amplitude is considered the A2 start, as seen on Figure 4. The P2 start instant is assumed to be the middle point detected used the improved Otsu method.

2.2 Hilbert transform IA and CWT

To improve the accuracy of the amplitude peak prediction of the S2 components, we implemented the algorithm presented by Mgdob et al. in [8]. The algorithm describes the application of a band-pass filter on the S2 sound signal, followed by the CWT. The scales chosen for the complex wavelet transform (CWT) should correspond to the frequencies ranging from 20Hz to 100Hz. To remove the irrelevant frequencies in the 20Hz to 200Hz region, a second order zero-phase Butterworth band-pass filter was applied.

The CWT was applied to the filtered signal using the Morlet wavelet. The chosen scales were from 300 to 3000, in increments of 300, to cover the frequency interval of 120Hz to 10Hz, characteristic of the A2 and P2 sound signals.

We register the frequencies with higher amplitude for each time instance, to approximate the instantaneous frequency of the signal. Frequency plateaus emerged at the time instance we believe to be the heart valves higher pressure points. To find these plateaus we

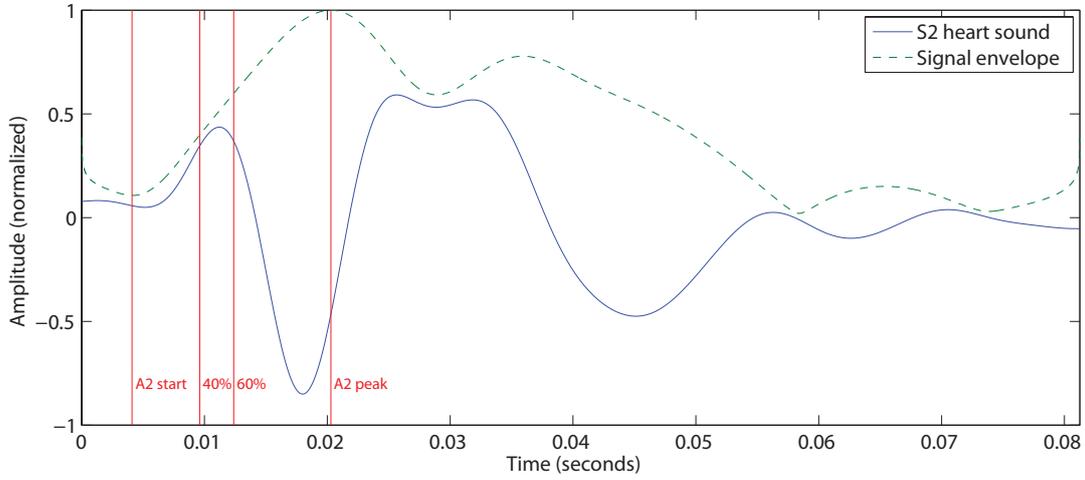


Figure 4: A2 sound start identification, including the 40% and 60% amplitude focal points. The solid line represents a S2 sounds signal, and the dashed line its envelope.

match the amplitude peaks found in the previous approach, based on the signal envelope, to the frequencies maximum within a 10ms region, as exemplified on Figure 5.

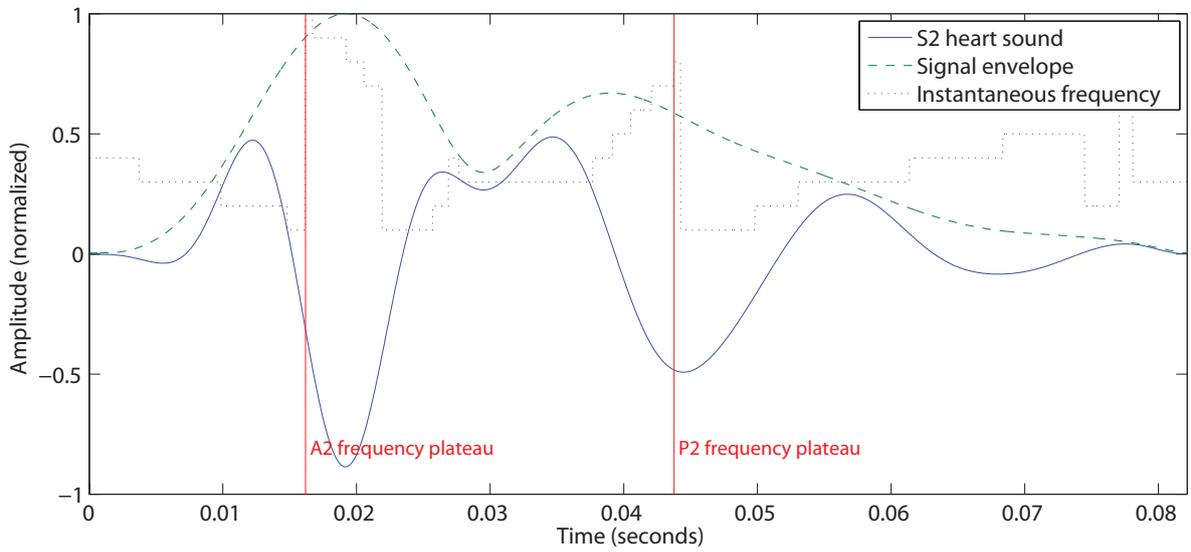


Figure 5: A2 and P2 frequency plateaus identification. The solid line represents the S2 sound signal envelope, the dashed line represents the signal envelope, and the dotted line represents the higher amplitude frequencies contour.

The maximum amplitude path presents different sections composed by amplitude lobes of variable duration for each sample. We predicted that both the A2 and P2 sound

durations should comprise the major lobes, one before and another after it. Therefore, we approximated the A2 start time instance to the valley before the first lobe before the A2 frequency identified earlier, as shows on Figure 6. The same with P2.

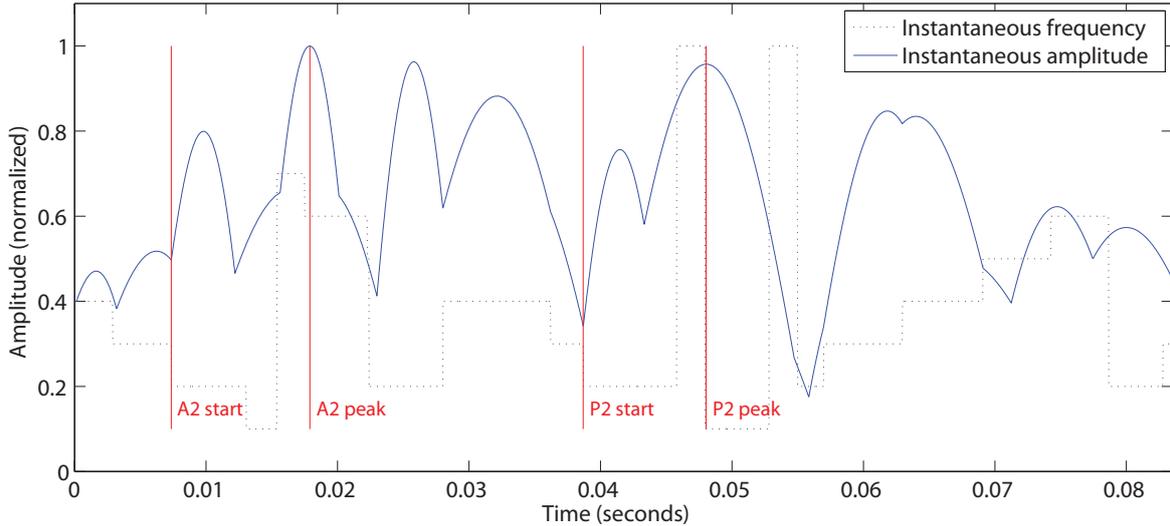


Figure 6: A2 and P2 amplitude peaks and start time instants identification. The dotted line represents the higher amplitude frequencies contour, and the dashed line the higher amplitudes, of the time-frequency values obtained from the CWT.

2.3 Hilbert transform IA and IF

The instantaneous frequency (IF) of the S2 sound signal was obtained again using the Hilbert transform. It corresponds to the derivative of phase of the Hilbert transform of S2. The resultant signal presents significant amplitude peaks where significant frequency changes occur in the original signal.

To facilitate the process of identifying the relevant amplitude peaks, the IF is post-processed. The signal is first filtered by a second order low-pass Butterworth filter. Then, its absolute value is retrieved. Finally, the signal is normalized by the third highest amplitude peak, to reduce the effect produced by occasional very high amplitude peaks. The last step is matching the correct IF amplitude peaks to the A2 and P2 start instants. The references used to delimit the search are the A2 and P2 IA amplitude peaks found using the approach described in section 2.1. It was observed that the component's start instant usually occur before the IA amplitude peaks. So, for each component, start instant is the instant of the highest IF amplitude peak between 50ms before and 10 ms after the component IA amplitude peak. Figure 7 presents the approach applied to a sample S2 sound signal.

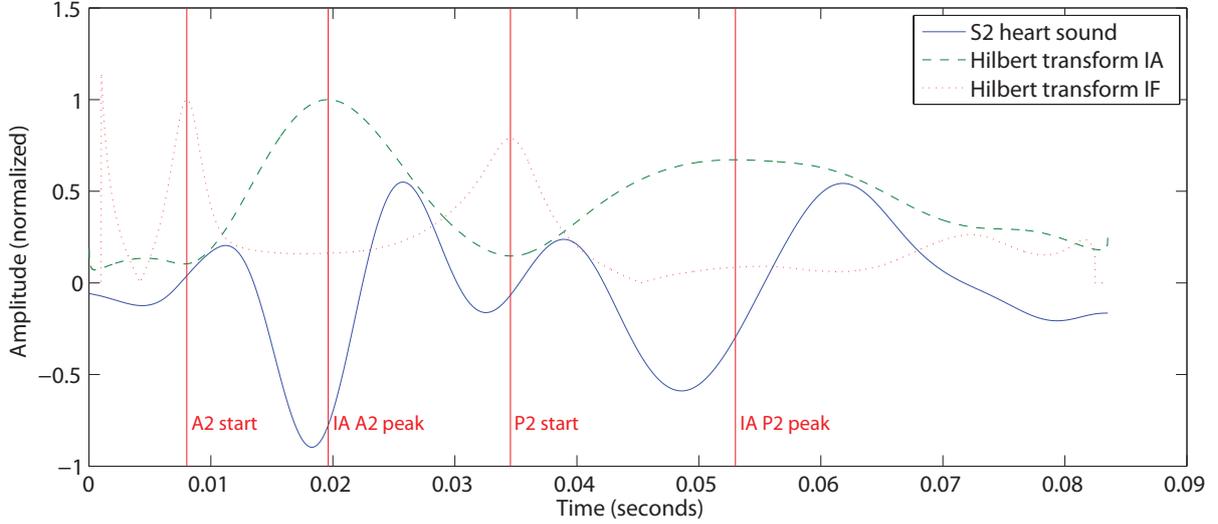


Figure 7: A2 and P2 start instants identification using the IF amplitude peaks. The solid line represents the S2 heart sound, the dashed line the IA, and the dotted line the post-processed IF.

3 RESULTS

In order to validate and compare the approaches developed for estimating the A2 and P2 start instants, 17 heart sound samples were acquired. The subjects belong to two groups. A first group of four healthy male volunteers with the following ages and body mass index (BMI) values: 42 years - $24,3kg/m^2$, 30 years - $31,4kg/m^2$, 24 years - $22,1kg/m^2$, and 22 years - $24,7kg/m^2$; and a second group of one volunteer with a mechanical aortic valve recently implanted, with unknown age and BMI. The results feature a total of 226 S2 heart sound samples, previously segmented. Table 2 summarized the results achieved by the methods described in section 2 in identifying A2 and P2 start instants.

	A2		P2	
	mean absolute error (ms)	standard deviation (ms)	mean absolute error (ms)	standard deviation (ms)
Method 2.1	5,0	3,5	13,9	18,1
Method 2.2	3,9	3,7	12,5	17,9
Method 2.3	3,5	2,9	10,3	11,3

Table 1: Comparison of the results of the three approaches for determining the start instant of the A2 and P2 components.

The Kolmogorov-Smirnov test was applied to assess the results' normality. The hypothesis of the results following a normal distribution could not be rejected, with a significance level of 5%. To evaluate if the difference between the methods' results was significant,

paired student t-tests were performed. Only the difference between the method 2.1 and 2.2, on the identification of the P2 start instant, could not be proven to be statistically significant, with a significance level of 5%. All other comparative results are determined relevant and can be compared through their mean values. Figure 8 presents the regression analysis made on the results and the correlation values calculated between the methods estimates and the initial annotations.

The method described in section 2.3 presents the best results. It exhibits a 30% improvement over method 2.1 and 10% improvement over method 2.2 on A2 start instant identification, and has a 26% improvement over method 2.1 and 18% improvement over method 2.2 on P2 start instant identification. Therefore, the combination of the IA and IF of the signal obtained through the Hilbert transform can be two important features to identify transition events in the S2 sound signal, specially the start of the A2 component. The identification of the P2 start instant of the component does not exhibit as good results as A2 in the outlined methods, which suggests that the methods presented do not account completely for wide variations of the S2 split duration. However, even when wide splitting is present, a peak of amplitude on the IF is observed on the P2 start instant, which reduces the challenge to the correct identification of the respective IF amplitude peak. A general classifier could be applied in this situation.

4 CONCLUSIONS

In this paper, three methods for the identifying the onsets of the A2 and P2, components of the S2 heart sound, were presented. The correct identification of these time instants leads to an accurate measure of the S2 Split, an important metric to preventively diagnose pulmonary hypertension. The methods were compared using an annotated database composed of 226 S2 heart sounds.

The obtained results suggest that the onsets of A2 and P2 can be adequately captured using features extracted from the signal's IA and IF. In fact, it was observed that the best onset identification results have been achieved with a new algorithm that combines these features. This seems to be consistent with other methods reported in literature where S2 is modeled with a combination of amplitude modulated chirp signals. In this context, the reported methods might be improved by incorporating this model explicitly. The proposed pattern recognition method does not explore a priori sequence information of sound components. These knowledge might be modeled using temporal pattern detection techniques such as hidden Markov models which might improve both robustness and accuracy of the method. It should be mentioned that the reported results should be considered preliminary since (i) a limited database of heart sounds has been applied, which (ii) have been collected mainly from a healthy population. Hence, another direction of further investigation will be the validation of the method using a more representative database collected from the typical target population.

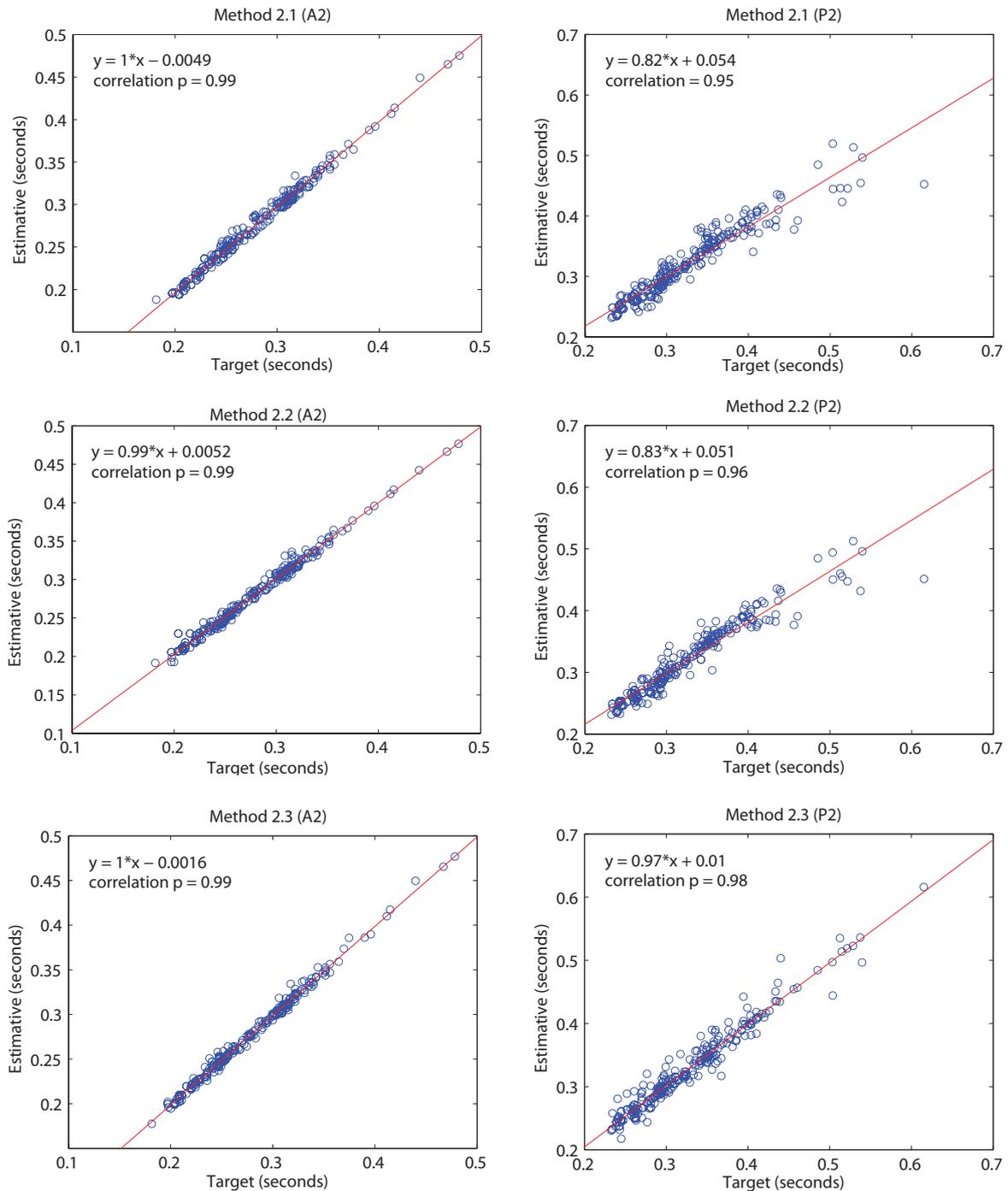


Figure 8: Regression analysis and correlation values for the identification of the A2 and P2 start instants using the three methods described.

REFERENCES

- [1] J. A. Shaver, R. Salerni, and P. S. Reddy, "Normal and abnormal heart sounds in cardiac diagnosis part i: Systolic sounds," *Current Problems in Cardiology*, vol. 10,

- no. 3, pp. 1 – 68, 1985.
- [2] T. D. Nauser and S. W. Stites, “Diagnosis and treatment of pulmonary hypertension,” *Am Fam Physician*, vol. 63, no. 9, pp. 1789–1799, 2001.
 - [3] H. Mgdob, J. Torry, R. Vincent, and B. Al-Naami, “Application of morlet transform wavelet in the detection of paradoxical splitting of the second heart sound,” in *Computers in Cardiology, 2003*, pp. 323–326, 2003.
 - [4] 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*, vol. 47, no. 10, pp. 1328–1335, 2000.
 - [5] V. Nigam and R. Priemer, “A dynamic method to estimate the time split between the a2 and p2 components of the s2 heart sound,” *Physiological Measurement*, vol. 27, no. 7, pp. 553–567, 2006.
 - [6] S. M. Debbal and F. Bereksi-Reguig, “Analysis and study of the variation of splitting in the second heartbeat sound of wavelet transform,” *Journal of Medical Engineering and Technology*, vol. 30, no. 5, pp. 298–305, 2006.
 - [7] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, pp. 62–66, Jan. 1979.
 - [8] I. R. Gray, “Paradoxical splitting of the second heart sound,” *British Heart Journal*, vol. 18, no. 1, pp. 21–28, 1956.