Real Time Identification of Heart Sounds Using Selectional Regional Correlation of the Time Frequency Domain

David Fourie
Department of Electrical
and Electronic Engineering
Stellenbosch University
Email: 16312023@sun.ac.za

M. J. (Thinus) Booysen
Department of Electrical
and Electronic Engineering
Stellenbosch University
Email: mjbooysen@sun.ac.za

***Abstract*—Cardiac auscultation is a non-invasive procedure that is mainly used in primary care, and involves diagnosis through analysis of the two heart sounds emanating from the cardiac cycle. None of the existing methods of computer aided auscultation can run in real time, and identify both S1 and S2 heart sounds, and operate without prior learning, and operate without making crude assumptions about the signal.**

This paper proposes a novel approach that is able to perform cardiac auscultation in real time and supports all of these features.

Existing approaches try to identify and employ unique features of S1 and S2, which could be different for different patients, equipment, placement of stethoscope, background noise, etc.

The proposed approach leverages the fact that distinct, dominant groups (e.g. heart sounds S1 and S2) will naturally emerge if the the time-frequency content of each sound is thresholded, correlated with the other sounds, and finally clustered.

The time-frequency information is derived from the continuous wavelet transform of the heart sound, and K-means is used for clustering.

The system was tested on a dataset of 230 recordings with over 5000 heart sound pairs, and test results show a predictive rate for both heart sounds of above 86 % – on par with existing approaches.

The system has been demonstrated working in real time, and an example application that uses this capability was developed.

I. Introduction

A. Background

Auscultation is a non-invasive method of diagnosis that medical doctors use to identify pathologies in the body. The process of auscultation involves listening to and analysing the sounds that emanate from organs. It is widely used in primary medical care as a low-cost method of diagnosis. Cardiac auscultation is the method of listening to the heart for diagnostic purposes. Beyond primary care, auscultation is replaced with other methods of diagnosis such as echocardiography, which uses sonar to visualise the movement of heart tissue and blood flow. Echocardiography, whilst considered the "gold standard" in diagnosis, is impractical for primary care because it requires the use of expensive equipment and specialist knowledge to interpret the results. Auscultation is therefore primarily used as a screening aid to determine whether a patient needs further examination.

978-1-4799-7560-0/15/$31 ©2015 IEEE

Computer aided heart auscultation is the process of using signal processing techniques and, in some cases machine learning, to automate the process of diagnoses of the heart through auscultation.

The cardiac cycle consists of a repeating set of contractions made by the four chambers of the heart. The chambers consist of the left and right atria and the left and right ventricles. There are four valves that control the movement of blood through the heart. The positions of these valves are shown in figure 1.

The path of blood through the heart explained below.

- The blood enters the heart through the vena cava and pulmonary vein. It fills the left and right atria.
- The atria contract and the blood enters the ventricles through the tricuspid and mitral valves.
- The ventricles contract and blood leaves the heart through the pulmonary and aortic valves into the pulmonary artery and aorta respectively.

The main sounds emanating from the heart can be classified as follows:

S1 - The "lub" of the "lub-dub" is termed the first heart sound. It is caused by the blood hammer as a result of the mitral and tricuspid valves closing.

S2 -The "dub" of the "lub-dub" is termed the second heart sound. It is caused by the blood hammer as a result of the pulmonary and aortic valves closing.

Murmurs - Turbulence in the blood flow in the heart causes a murmur to occur. They are either physiological or pathological in nature.

The time between the closing of the valves in the second heart sound is of clinical significance. When the valves close at different times, the heart sound is said to be split.The duration of this split and how it varies with breathing can be used to identify certain conditions in the heart. There are four main types of S2 split.Normal splitting, where the split increases with inhalation and decreases with exhalation; wide splitting, when the split is identified to be wider than normal splitting; fixed splitting, where the split remains constant during inhalation and exhalation; and paradoxical splitting, where the split increases with exhalation and decreases with inhalation. The abnormal splitting of the second heart sound is used as a method of diagnosis for certain heart conditions [1].

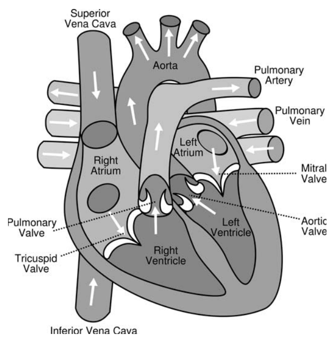


Fig. 1. Diagram showing the movement of blood through the heart

A key challenge in automated cardiac auscultation is heart sound segmentation, which refers to the identification of the physiological features of the different heart sounds to determine their relative physiological timing. Several techniques have been developed to detect the timing of the sounds, using peaks in the energy envelope [2], high frequency signatures in wavelet coefficients [3], neural networks [4], empirical mode decomposition [5], Hidden Markov Models [6] and selectional regional correlation [7].

B. Contribution

The ideal Heart Sound Segmentation system should be able to accurately identify S1 and S2, use as few assumptions about the signal as possible to reduce the effect of noise, should avoid the use of prior learning to increase generality across patients and conditions, and should be able to work on a small segment of a signal so that it can be used in a close to real time environment. The system proposed in this paper is able to satisfy these requirements.

This paper builds on the work of [7] by using the correlation of multiple templates to identify a signal segment as opposed to a single template. This method is used to implement a system that can identify the two main heart sounds in close to real time. An example application is developed that uses this method to identify S1 and S2, and to detect the split in the second heart sound.

Table I provides a summary of the related work and the contribution of this paper.

Property / paper	[4]	[5]	[3]	[6]	[7]	This paper
S1 and S2	×	✓	✓	✓	✓	✓
Few assumptions	✓	✓	×	×	✓	✓
Adaptive learning	×	✓	✓	×	×	✓
Works in real time	×	×	×	×	×	✓

TABLE I
TABLE HIGHLIGHTING THE CONTRIBUTION OF THIS WORK AGAINST THE STATE OF THE ART.

II. MATHEMATICAL BACKGROUND

In this section, an overview of the mathematical tools required for the used to develop the algorithm.

A. Time frequency analysis

The signals that emanate from the heart have frequency content which varies with time [8], therefore Fourier analysis of the signals provide limited information. Time frequency analysis examines how the frequency content of a signal changes with time and therefore will yield useful information for analysis. The time-frequency content of the signal is extracted using wavelet decomposition. Wavelet analysis identifies the different frequency components at different resolutions. The equivalent Fourier transform method is the Short Time Fourier Transform, which performs the analysis at a fixed resolution that is dependent on the window size.

Using the variable names from [4], the wavelet transform of a function f at a time u and scale s is:

$$W_f(u,s) = \int_{-\infty}^{\infty} f(t)\Psi^*_{u,s}(t)dt \tag{1}$$

where

$$\Psi_{u,s}(t) = \frac{1}{\sqrt{s}}\Psi\left(\frac{t-u}{s}\right) \tag{2}$$

u and s represent the time delay and frequency scale variables. $\Psi(t)$ called the mother wavelet as it generates the different wavelets required for the multiple levels of resolution. The mother wavelet used in this paper is the Morlet wavelet.

$$\Psi(t) = \pi^{-\frac{1}{4}}\left(e^{-j\omega_0 t} - e^{-\omega_0^2/2}\right)e^{-t^2/2} \tag{3}$$

B. Clustering

K-means clustering is a statistical method of grouping together a set of vectors into a specified k number of clusters [9]. Each cluster is associated with a mean which determines how the date points are assigned to each cluster. The algorithm iterates over the following steps until the assignment of vectors does not change between iterations.

Assignment Each vector is assigned a cluster based on minimising the Euclidean distance from each particular vector to the centroid of the cluster.

Update New centroids calculate by finding the mean of all the vectors assigned to a particular cluster,

III. METHOD

An overview of the system can be seen in figure 2.

A. Development of S1 and S2 identification

1) Heart sound location identification: To identify the origins of all the sounds in a signal, the location of each sound is required. This is achieved by identifying the peaks in the Shannon envelope of the signal. The Shannon envelope

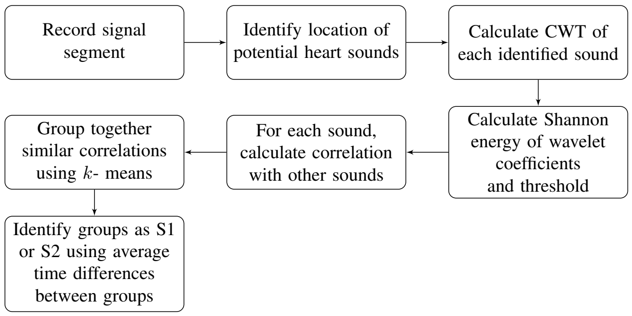


Fig. 2. Flow diagram showing the layout of the S1 - S2 detection system

is a transformation of the signal that emphasises the medium amplitude values of the signal de-emphasising high and low amplitudes. The resulting transformation is a smooth, continuous function in which the peaks represent the location of the sounds. The Shannon envelope of a heart sound segment is shown in figure 3a. The calculation for the Shannon envelope is shown below:

$$S(t) = \frac{1}{10} \sum_{i=-5}^{i=5} x(t+i)^2 \times \log\left(|x(t+i)|\right) \tag{4}$$

where $x(t)$ is the signal of which the Shannon envelope $S(t)$ is calculated.

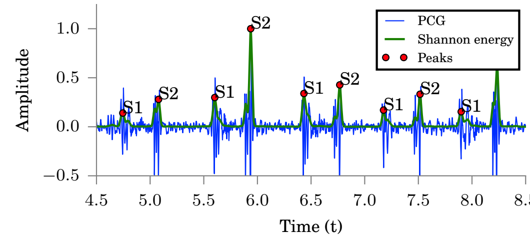
2) Time frequency data extraction: The CWT for each sound is calculated from a segment of signal around each located peak. The segment length is calculated from the width of the peaks in the Shannon envelope. The magnitude of the CWT of the signal in figure 3a is shown in figure 3b.

3) Time frequency transformation: Correlation of the time frequency components is used to group together similar sounds. To make this process more accurate the following properties are needed of the signals to be correlated:

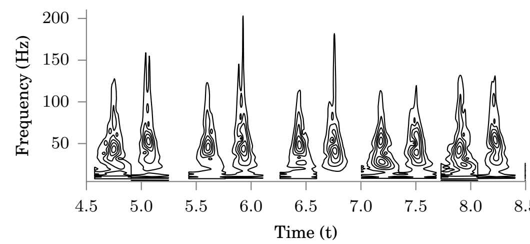
- The peak amplitude of the signals needs to be constant, otherwise higher valued signals bias the correlation
- The effect of noise needs to be minimised

To achieve these properties, the Shannon energy envelope is calculated for the coefficients. The Shannon energy is used for similar reasons as it was in the envelope. It reduces the effect of high and low amplitude noise. The transformed coefficients are then thresholded to values of 1 or 0 using high and low thresholds. This is to eliminate the effect of amplitude on the correlation as well as emphasising a specific shape and position for the sound in the domain. The result of the transformation and correlation is shown in figure 3c.

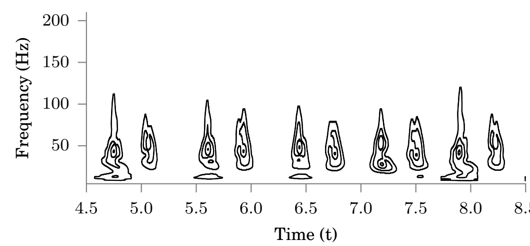
4) Selectional regional correlation as a feature vector: To segment the located sounds, the selectional regional correlation for each sound is calculated with respect to all the other sounds in the signal. The time frequency correlation for two sounds is calculated by multiplying the components of the two sounds together and summing the result. The correlation for a particular sound are expected to increase when the compared sound is similar and decrease when the sound is dissimilar. The patterns of increasing and decreasing levels of similarity are shown in figure 3d.



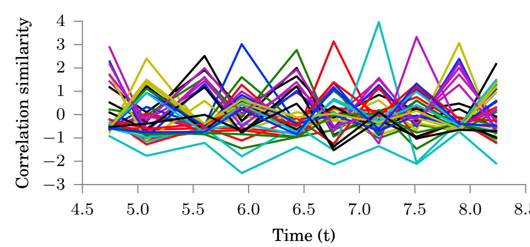
(a) Graph showing the Envelope, PCG and peaks



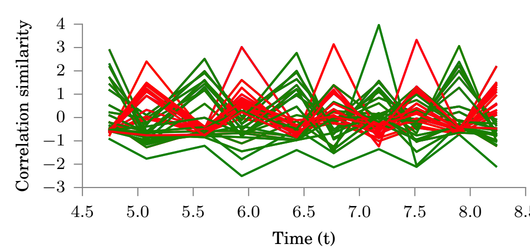
(b) Graph showing wavelet coefficients of the signal



(c) Graph showing transformed wavelet coefficients of the signal



(d) Graph showing the correlation of the transformed wavelet coefficients



(e) Graph showing the grouped correlations of the signal

Fig. 3. Figure showing the development of the system using a sample of a PCG

5) Clustering the heart sounds: The selectional regional correlation for a particular sound is used as a feature vector for a machine learning system. The features in a particular vector are its similarity with the sounds. The heart sounds are then clustered using the k-means clustering algorithm. The clustered correlations are shown in figure 3e

6) Identification of clustered heart sounds: The clustered correlations are then used to separate the heart sounds into two groups. These groups will each contain a number of sounds that are similar to each other. The next step is to identify the source of each group of similar sounds. The time difference between successive S1 and S2 sounds is smaller than time difference between successive S2 and S1 sounds. Using this information, the groups are identified as either S1 or S2.

B. Real time analysis

An overview of the real time system is shown in figure 4.

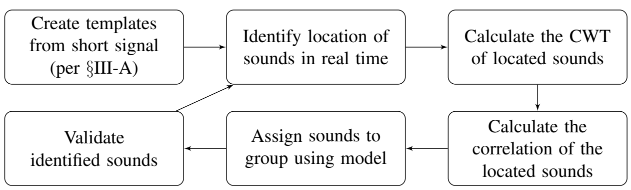


Fig. 4. Flow diagram showing the layout of the real time system

1) Model creation: To identify a heart sound in real time, the system first needs to generate templates of S1 and S2 sounds with which an unknown sound can be correlated. The templates are constructed from the transformed time frequency coefficients of identified S1 and S2 an intitial recording (20 seconds in this case). The templates are arranged in an alternating pattern of groups, namely S1 and S2. With the templates in a pattern, the correlation of the templates is calculated and the mean value of the correlation in each group is found. This is the equivalent of performing the k-means algorithm on the correlations.

2) Identify location of sound: The real time analysis of the signal is performed on overlapping 2 second chunks of data with a refresh rate of 1 second. These intervals guarantee that at least one sound is captured in the chunk as the average heart rate is 60 to 100 beats per minute. In each 2 second chunk of data, the Shannon envelope is calculated and the peaks are found. The peaks represent the location of a potential heart sound.

3) Calculate time frequency content of located sounds: The time frequency content of each identified sound is calculated and transformed in the same way as subsections III-A2 through III-A3.

4) Calculate correlation of identified sounds: The identified sounds are correlated with the S1-S2 template to generate feature vectors to identify the sounds.

5) Identification of sounds: The sounds are identified by calculating the Euclidean distance from their feature vectors to each of the means calculated in section III-B1. The group with the smallest distance is chosen as the label for the sound being identified.

6) Validation of identified sounds: To increase the accuracy of the system, a simple method of validation is used. The method consists of only using pairs of sounds that correspond to the physiological process. Therefore the only sound pairs that are used are the S1 to S2 pair and the S2 to S1 pair. Any other pair is considered to be an error in identification.

IV. Results

A. S1 and S2 detection

The system was tested using a database of 230 heart sound signals recorded from various areas of the heart. The signals were selected from a larger database that was recorded at the Red Cross Children's hospital under ethical conditions using the SensiCardiac© system. To ensure a fair test of the system, a set of visual criteria was used to select suitable signals for testing. The criteria are:

- Visible S1 and S2
- Few extra sounds in recording
- Minimal signal clipping

Each signal was recorded over a 15 second period using an electronic stethoscope. The signals were recorded at 22 kHz. To test the system, the heart sounds in each signal were tagged manually as S1, S2 or other, using an ECG to assist. The average processing time for the signals was 1 second using a PC with an Intel Core i5 and 4 GB of RAM.

The results for the system are presented in table II which shows how the different heart sounds were classified.

	Manually identified S1	Manually identified S2	Manually identified other
System identified S1	2444	508	15
System identified S2	334	2168	18

TABLE II
TABLE TO SHOW RESULTS OF THE SYSTEM IMPLEMENTED IN THIS PAPER

Table III shows the results for a system implemented using one of the methods in [2].

	Manually identified S1	Manually identified S2	Manually identified other
System identified S1	1919	290	6
System identified S2	325	1682	8

TABLE III
TABLE TO SHOW RESULTS FOR A SYSTEM IMPLEMENTED USING ONE OF THE METHODS IN [2]

The tables show that the system implemented in this paper has an accuracy of 86% with an S1 predictive value of 87% and an S2 predictive value of 85%. The alternative system fares slightly better with an accuracy of 89% with an S1 predictive value of 88% and an S2 predictive value of 89%. However the system described in this paper correctly identified 27% more S1 sounds and 29 % more S2 sounds, and works in real time.

B. Real time results

The real time system was implemented using the S1 and S2 detection method described in this paper to generate the templates. A Thinklabs One stethoscope was used to record the data. Using a sample rate of 1 kHz a computer with an i5 processor was able to identify the origin of the sounds within one second of the sounds being recorded. A video of the real-time detection can be seen at [10]. The video shows a 60 second real time analysis, where the heart sounds are being detected. In this part of the video there were 39 pairs of S1 and S2 sound, of these sounds 24 were identified by the computer as S1 and S2 pairs, with no incorrectly identified sounds.

V. Conclusion

This paper has demonstrated how the heart sounds in a signal can be identified in real time using selectional regional time frequency correlation. The locations of possible heart sounds were identified using the peaks in the Shannon energy of the signal. The time-frequency characteristics for each identified heart sound were calculated. The time-frequency characteristics were then transformed using Shannon energy. The time-frequency characteristics for each heart sound were correlated with all the other heart sounds in the signal. The correlations corresponding to each sound were used as a feature vector for clustering using K-means. The correlations were clustered into two groups. The groups were identified using the physiological timing differences between S1 and S2. A database of 230 heart sound signals was used to test the offline system. The testing showed that the system implemented in the paper was as accurate as well being sensitive to the detection of S1 and S2. The system was then tested in real time and an example application of detecting the second heart sound was provided.

References

[1] J. M. Felner, "The Second Heart Sound," in *Clinical Methods: The History, Physical, and Laboratory Examinations.*, 3rd ed., H. Walker, Hall WD, and Hurst JW, Eds. Butterworths, 1990, ch. 23. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK341/

[2] S. Choi and Z. Jiang, "Comparison of envelope extraction algorithms for cardiac sound signal segmentation," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1056–1069, Feb. 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0957417406003927

[3] D. Kumar, P. Carvalho, M. Antunes, J. Henriques, L. Eugenio, R. Schmidt, and J. Habetha, "Detection of S1 and S2 heart sounds by high frequency signatures." *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 1, pp. 1410–6, Jan. 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/17946890

[4] T. Oskiper and R. Watrous, "Detection of the first heart sound using a time-delay neural network," in *Computers in Cardiology*. IEEE, 2002, pp. 537–540. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1166828

[5] C. D. Papadaniil and L. J. Hadjileontiadis, "Efficient Heart Sound Segmentation and Extraction Using Ensemble Empirical Mode Decomposition and Kurtosis Features," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1138–1152, Jul. 2014.

[6] A. D. Ricke, R. J. Povinelli, M. T. Johnson, and G. E. Healthcare, "Automatic Segmentation of Heart Sound Signals Using Hidden Markov Models," pp. 953–956, 2005.

[7] E. Sejdić and J. Jiang, "Pattern Recognition in Time-Frequency Domain: Selective Regional Correlation and Its Applications," in *Pattern Recognition Techniques, Technology and Applications*, Y. Peng-Yeng, Ed. Vienna: I-Tech, 2008, ch. 27, pp. 13–626.

[8] S. M. Debbal and F. Bereksi-Reguig, "Analysis of the second cardiac sound using the Fast Fourier and the Continuous Wavelet Transforms," *Internet Journal of Medical Technology*, vol. 3, no. 1, 2006.

[9] J. A. Hartigan and M. A. Wong, *Applied Statistics*, vol. 28, no. 1, p. 100, Jan. 1979.

[10] D. Fourie, "Real time segmentation of heart sounds," 2015. [Online]. Available: https://www.youtube.com/watch?v=W9M6rh134p4