TECHNICAL PAPER

Autonomous detection of heart sound abnormalities using an auscultation jacket

C. Visagie[1], C. Scheffer[1], W.W. Lubbe[2] and A.F. Doubell[2]

[1]*Department of Mechanical and Mechatronic Engineering, Stellenbosch University, Matieland, South Africa*
[2]*Cardiology Unit, Department of Internal Medicine, Faculty of Health Sciences, Stellenbosch University, Tygerberg, South Africa*

Abstract

This paper presents a study using an auscultation jacket with embedded electronic stethoscopes, and a software classification system capable of differentiating between normal and certain auscultatory abnormalities. The aim of the study is to demonstrate the potential of such a system for semi-automated diagnosis for underserved locations, for instance in rural areas or in developing countries where patients far outnumber the available medical personnel. Using an "auscultation jacket", synchronous data was recorded at multiple chest locations on 31 healthy volunteers and 21 patients with heart pathologies. Electrocardiograms (ECGs) were also recorded simultaneously with phonocardiographic data. Features related to heart pathologies were extracted from the signals and used as input to a feed-forward artificial neural network. The system is able to classify between normal and certain abnormal heart sounds with a sensitivity of 84% and a specificity of 86%. Though the number of training and testing samples presented are limited, the system performed well in differentiating between normal and abnormal heart sounds in the given database of available recordings. The results of this study demonstrate the potential of such a system to be used as a fast and cost-effective screening tool for heart pathologies.

Key words auscultation, neural network, electronic stethoscope, heart sound

Introduction

According to the World Health Organization estimates of 2003, cardiovascular disease (CVD) accounts for approximately 16.7 million deaths globally, which amounts to over 29% of all deaths globally[1]. The Medical Research Council of South Africa lists cardiovascular disease as second only to AIDS as the cause of death in South Africa, accounting for 16.6% of all deaths[2]. In a report produced by Columbia University on the impact of cardiovascular disease in developing countries, the mortality rate in South Africa due to cardiovascular disease is listed as 199 per 100 000 people with a total mortality rate of 481 per 100 000 people[3]. Although no accurate figures for valve diseases are available for South Africa, the prevalence is believed to be higher than in the developed world[4].

The impact of the burden of cardiovascular disease is thus obvious. Less obvious, is the limited ability of regular healthcare personnel to detect cardiovascular disease, especially in the developing world and in rural African communities. In most cases, clinicians will have to rely on the mechanical stethoscope for an initial diagnosis, due to the lack of specialized technologies and skills, and available time for one clinician to examine so many patients. However, cardiac auscultation is a difficult clinical skill to master, and human hearing has limitations with respect to cardiac auscultation[5]. Most medical practitioners do not feel confident that they will be able to identify abnormal heart sounds, and this places the patient and the medical practitioners at risk[5,6]. The reliance on new cardiac imaging technologies such as echocardiography, computed tomography and magnetic resonance myocardial imaging may have placed this skill under further threat[6]. Additionally, sophisticated imaging modalities are not always accessible to patients, particularly in the developing world, and the need to analyze the diagnostic information contained in heart sounds remain a primary tool for initial diagnosis.

Normal heart sounds are the sounds generated when heart valves close. Abnormal heart sounds are generated as the blood flows though an abnormally small valve, or blood flowing through a valve that is not closing properly. Depending on the location of the abnormal heart sound, as well as the pitch, the variations in amplitude, and the duration of the sound, an experienced practitioner can determine which valves are affected. Many heart sounds that may sound abnormal to the human ear, are often physiologically normal, and hence will not require

Corresponding author: C. Scheffer, Department of Mechanical and Mechatronic Engineering, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa
Tel: +27 21 808 4249, Email: cscheffer@sun.ac.za
Received: 19 August 2009; Accepted: 19 October 2009

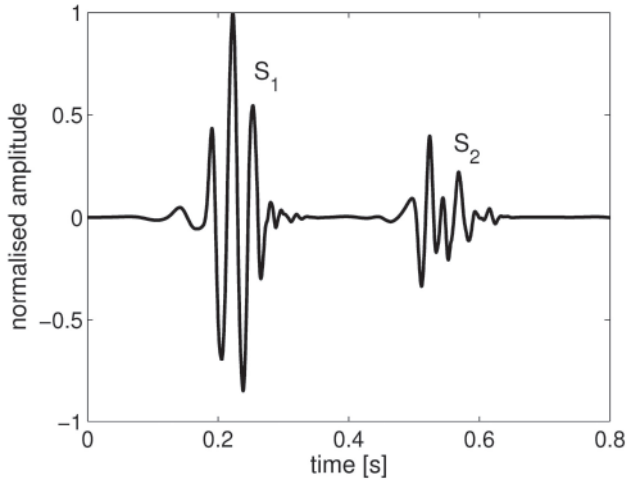

Figure 1. *Normal heart cycle with S_1 and S_2 indicated.*

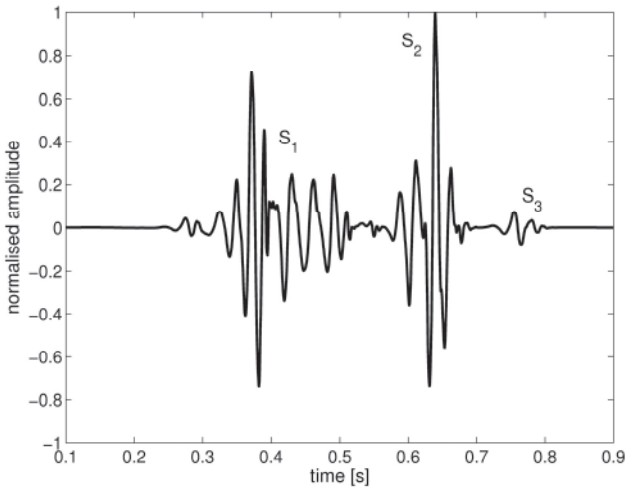

Figure 2. *Abnormal heart cycle (mitral regurgitation) with S_1, S_2 and S_3 indicated.*

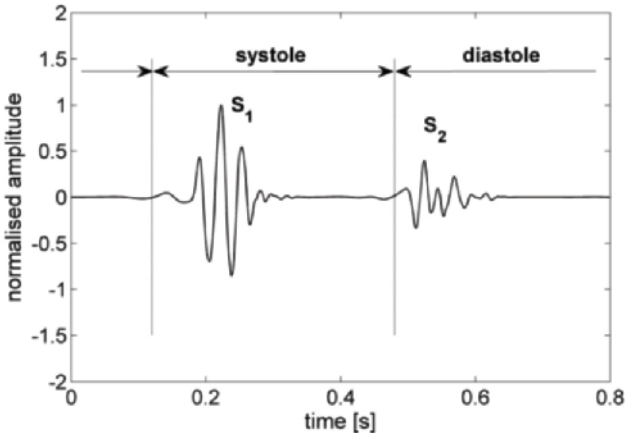

Figure 3. *Systole and diastole of normal heart sound.*

specialist investigation or intervention. Hence good auscultation and interpretation of heart sounds can potentially limit unnecessary specialist referrals.

Currently, electronic stethoscopes only record in a single position at a time, and cannot record all the heart sounds at the different positions on the chest wall simultaneously. These sounds are also typically not synchronized with an electrocardiogram (ECG) although algorithms have been proposed that segment phonocardiograms without the use of an ECG[7]. In this study, an "auscultation jacket" and a heart sound classification system are presented, which are aimed specifically to improve healthcare delivery in rural and medically underserved communities. The auscultation jacket records heart sounds at multiple locations on the chest simultaneously, and also synchronizes these recordings with an ECG. After data processing, the system can aid in classifying the heart sounds as "normal" or "abnormal".

The overall aim of this work was to demonstrate the potential of the proposed system to facilitate screening of patients for heart pathologies, especially in the developing world. It is envisaged that the system can eventually be used by a lesser skilled healthcare worker to record data that can be sent to a doctor via a telemedicine connection for review.

Theoretical background

Heart sounds

In the heart, a complete cycle for pumping blood consists of two stages, namely systole and diastole. Systole is the contraction of heart chambers, driving blood out of the chambers. Diastole is the period of time when the heart fills with blood after systole (contraction). Ventricular diastole is the period during which the ventricles are relaxing, while atrial diastole is the period during which the atria are relaxing. All four chambers of the heart undergo systole and diastole in a timed fashion so that blood is propelled forward through the cardiovascular system.

During a complete cycle, four general heart sounds, known as S_1, S_2, S_3 and S_4 are formed. S_1 and S_2 are the heart sounds that one normally associates with the beating heart ("lubb-dupp"). According to Rangayyan[8], S_1 is produced by four different components that range from the closure of the mitral and tricuspid valves to the deceleration of the pumped blood and turbulence between the blood and the aorta. S_2 is produced by the closure of the aortic and pulmonary valves. The third heart sound (S_3) can also sometimes be heard and is due to the sudden termination of the ventricular rapid-filling phase. The fourth heart sound (S_4) occurs at the same time as (and is due to) atrial systole. S_4 is usually a low-pitched sound and best heard at the apex of the heart. The period from the onset of S_1 to the onset of S_2 is associated with systole, and the period from the onset of S_2 until the start of the next cycle (S_1) is associated with diastole. One normal heart sound cycle is shown in Figure 1, with S_1 and S_2 indicated, and one abnormal heart cycle of a patient that suffers from mitral regurgitation, is shown in Figure 2. The relationship between the two main heart sounds (S_1 and S_2) and systole and diastole, is indicated in Figure 3.

Table 1. *Patient diagnoses.*

Patient	Diagnosis
1	AS; MS; AR; MR; SBE on AV
2	MS; MR; MMVD
3	MR
4	VSD ; PR
5	SBE on AV; AI
6	SBE on AV; AR
7	AR
8	AS; AR; MS
9	VSD
10	MR; Valve Lesions
11	AR; AS
12	MR; PR; AR; MR; SBE on MV
13	AR; MR; SBE on AV
14	AR; MS; MR
15	AS; AR; MR; MAVD
16	MR; SBE on MV
17	VSD
18	MS; MR; MAVD
19	MR
20	AR
21	MAVD; MMVD
Abbreviations	
AS	Aortic Stenosis
AR	Aortic Regurgitation
MS	Mitral Stenosis
MR	Mitral Regurgitation
PR	Pulmonary Regurgitation
VSD	Ventricular Septal Defect
AI	Aortic Insufficiency
MMVD	Myxomatous Mitral Valve Disease
SBE	Subacute Bacterial Endocarditis
MAVD	Mitral and Aortic Valve Disease
AV	Aortic Valve
PV	Pulmunic Valve
MV	Mitral Valve
HS	Heart Sound

Classification of heart sounds

In this work, several signal processing techniques are combined with an Artificial Neural Network (ANN) for the autonomous classification of the sounds into two categories, namely "normal" and "abnormal". Previous authors have also addressed the analysis and classification of heart sounds[9-13]. Signal processing techniques implemented in the analysis of heart sounds include the Fast Fourier Transform (FFT), Short-Time Fourier Transform (STFT), Wigner Distribution (WD), Choi-Williams Distribution (CWD) and the Wavelet Transform (WT).

Obaidat[5] concluded that the WT performs best in identifying the components of S_1 and S_2 when compared to the STFT and WD. Bentley et al.[14] evaluated the STFT, WD, CWD and the WT as techniques to extract information from phonocardiograms. According to Bentley et al.[14], the CWD has superior resolution compared to the other techniques, but the WT feature extraction technique performed the best in classifying native and bioprosthetic heart sounds.

Ölmez and Dokur[15] implemented the Daubechies wavelet of order two (db2) in analyzing their signals. The second level detail coefficients were broken up into 32 subwindows of 128 samples each and the power of each subwindow calculated. The resulting power curve was then implemented as input to their grow-and-learn (GAL) network classification scheme, obtaining a correct classification rate of 99.4%. Gupta et al.[13] used homomorphic filtering to identify S_1 and S_2, and then used the same approach as Ölmez and Dokur to construct and train their neural network. Cathers[10] used the amplitude envelope of the recorded heart sounds as input to a feed-forward neural network, obtaining a classification rate of 100%.

DeGroff et al.[11] developed an ANN to differentiate between innocent and pathological murmurs in children. A feed-forward network trained with the back-propagation algorithm was implemented. The input space comprised the magnitude of the Fourier coefficients in the 0-150 Hz range. A sensitivity of 100% and a specificity of 100% were obtained. Andrisevic et al.[9] implemented an ANN consisting of two hidden layers and one output layer to differentiate between normal and abnormal heart sounds. The network was trained with the back-propagation algorithm and a sensitivity of 64.7% and a specificity of 70.5% was obtained. In this paper, the approach differs from previous work in that heart sounds are recorded at several locations on the chest simultaneously and synchronized with an ECG. For this study, signal processing methods found in the literature were employed with some modifications.

Methodology

The heart sounds and ECGs of 31 healthy volunteers and 21 patients who suffer from valve-related or auscultatory abnormalities, were recorded. The experimental protocol for the study was approved by the committee for human research at Stellenbosch University, South Africa, in accordance with international practices and all participants gave informed consent. All the participants were individually assessed by a cardiologist at Tygerberg Academic Hospital in Cape Town, South Africa. The cardiologist used a 12-lead ECG and an echocardiogram to obtain a diagnosis for all the participants (including the healthy participants). The diagnoses of the patients with heart abnormalities are given in Table 1. All the participants were recorded in the supine position, breathing normally. The duration of the recordings was 10 seconds.

Figure 4. *Anterior and lateral parts of the jacket that demonstrates all the ECG leads and multiple stethoscopes.*

Figure 5. *Posterior chest wall part of the jacket to record lung sounds.*

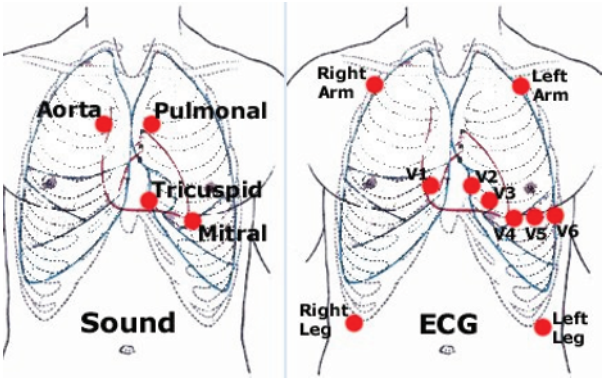

Figure 6. *Positions for sound and ECG recordings.*

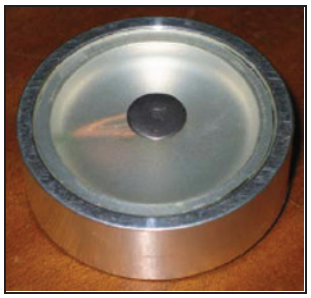

Figure 7. *ECG lead built into the diaphragm of an electronic stethoscope.*

Hardware

To facilitate multiple simultaneous recordings, electronic stethoscopes and ECG electrodes for a 12-lead ECG were built into an auscultation jacket, as was previously reported by Koekemoer and Scheffer[14]. The jacket consists of neoprene sheets and adjustable straps, with stethoscopes and ECG leads embedded at specific locations. The jacket was designed as a prototype to fit adult men with an anthropometric size in the 50th percentile of the South African male population.

The jacket has 21 electronic stethoscopes built into the anterior, lateral (Figure 4) and posterior (Figure 5) parts. The posterior stethoscopes were used to also record the lung sounds and hosted the two limb electrodes of the left and right arm. A "neck collar" is included that has two electronic stethoscopes to record Carotid bruit sounds, though the lung sound recordings were not used for the purpose of this study.

The jacket was designed in such a manner that all sensors would be positioned at constant locations on the chest and neck of a patient. The heart sounds of the different heart valves and murmurs, if present, can maximally be heard over specific positions on the precordium of a patient, and hence these locations were chosen for the stethoscopes. Six precordial leads of a 12-lead ECG overlaps with the auscultation points for the tricuspid and mitral valves as indicated in Figure 6. For these cases, the ECG electrode was built into the diaphragm of the electronic stethoscope (Figure 7).

The stethoscopes are positioned at the normal auscultatory positions, i.e. the second left and right intercostal space, fourth left and right intercostal space, fifth left intercostal space and sixth left intercostal space. The ECG electrodes were positioned at the normal positions for V_1-V_6, the left arm and right arm electrodes were placed on the left and right shoulders respectively, while the left and right leg electrodes were placed on the left and right hips. All the hardware was customized specifically for the auscultation jacket and fabricated in-house. The sensors communicate via USB with a host computer, and a customized powered USB hub was constructed to enable all the sensors to communicate in sync with the host computer. The recordings were done with a sample rate of 2 kHz on all channels. The complete auscultation jacket is shown in Figure 8, and further developments of this hardware is also described by Koekemoer and Scheffer[14].

After recording the data using the jacket, these steps were followed: denoising, segmentation (i.e. identify systole and diastole in each heart sound cycle), feature extraction (i.e. extract signal features indicative of an abnormal heart sound) and ANN training and testing. These steps are described to more detail in the sections that follow.

Denoising

Due to the presence of unwanted noise and interference, the recorded signals had to undergo a denoising procedure. The ECG signals were filtered with a low-pass fourth order Butterworth filter (f_c = 40 Hz),

Figure 8. *Jacket fitted over a mannequin, demonstrating an anterior, lateral and posterior view.*

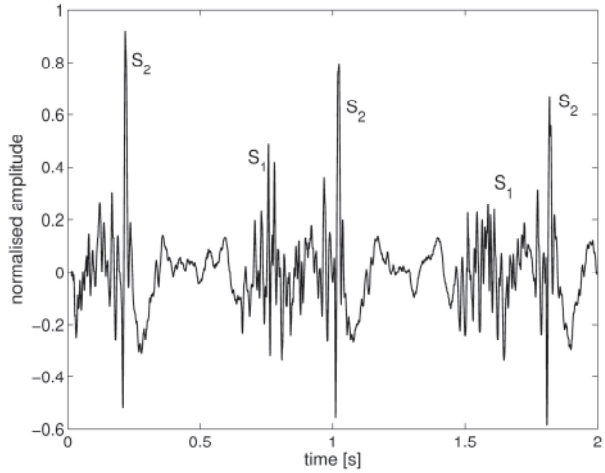

Figure 9. *Abnormal phonocardiogram of patient suffering from mitral regurgitation.*

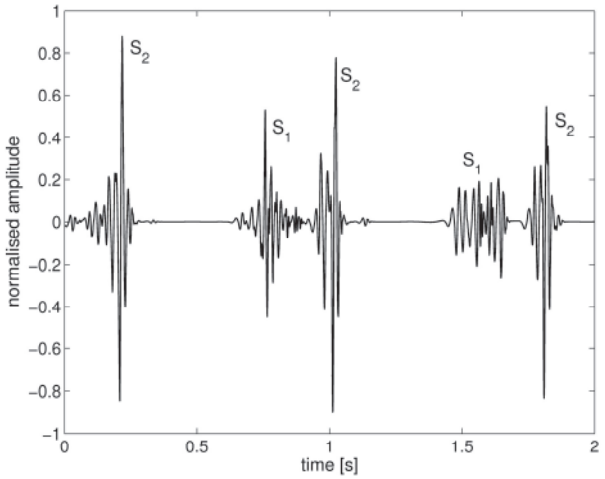

Figure 10. *Denoised phonocardiogram (db7 level 5 wavelet denoising) of patient suffering from mitral regurgitation.*

ensuring that potential interference from the main electrical supply was filtered out.

Denoising was performed by wavelet soft threshold denoising, similar to previous authors[12,16]. Only orthogonal wavelets were considered for denoising purposes as this allows perfect reconstruction of the denoised signal[15]. Messer *et al.*[15] studied the effects of different wavelets on denoising recorded heart sounds and found that certain wavelets from the Haar, Coiflet, Daubechies and Symlet families provided the best results. Also according to Messer *et al.*[15] the more a wavelet resembles the signal, the better it denoises the signal. The Daubechies wavelet was selected due to the unsymmetrical nature of this specific wavelet (heart sounds are also unsymmetrical in nature) and the close resemblance to recorded heart signals. The Haar and Coiflet family of wavelets were deemed too symmetrical to be used for denoising purposes and the Symlet family of wavelet proved to be too computationally intensive for denoising purposes as shown by Messer *et al.*[15] According to Messer *et al.*[15] a decomposition level of 5 produced reasonable results while decomposing the signal further often produced marginal benefits and increases computation time. The authors of this study also found that increasing the decomposition level above 5 did not produce significantly better results. The Daubechies 7 wavelet at a decomposition level of 5 was experimentally determined to provide the best results. Although Messer *et al.*[15] suggested that Daubechies wavelets of order 11, 14 and 20 provided better results the authors of this study found that wavelets of an higher order did not produce better denoising results and in some instances provided worse results. Also, a single wavelet for denoising purposes were desired to automate the process as far as possible and the above mentioned Daubechies wavelet proved to provide the best results. Figure 9 shows an original recorded heart sound signal, and Figure 10 shows the signal after the denoising procedure.

Two wavelet thresholds were calculated and the larger of the two was used. The first threshold was calculated by using Stein's Unbiased Estimate of Risk (SURE)[11]. This threshold was multiplied by 0.3 in order to reduce the threshold since a larger threshold resulted in S_1 or S_2 being discarded. The second threshold value was calculated by first calculating the standard deviation of the first-level detail of the DWT. The assumption is that the majority of the noise in the signal is captured in this level[17]. The standard deviation is then multiplied by four to yield the second threshold.

Segmentation

The next step in the process was to identify the heart sound cycles in the recordings and split them into systole and diastole. Segmentation algorithms exist that do not utilize the ECG signal[19]. The authors of this study had the benefit of access to the ECG signal and therefore decided to implement it in segmenting the phonocardiogram[1].

Three cycles of each recorded heart sound were extracted by identifying the QRS-peaks in the ECG. The QRS-peak can be taken as the start of a cycle and hence the onset of S_1 [8]. To accentuate the QRS-peak and to attenuate the P-wave and T-wave in the ECG, a QRS-peak identification algorithm described by Rangayyan[8] was

[1]*The aim of this study was not to develop a segmentation algorithm or determine which is algorithm is superior to others and therefore no conclusion on this can be made in this study.*

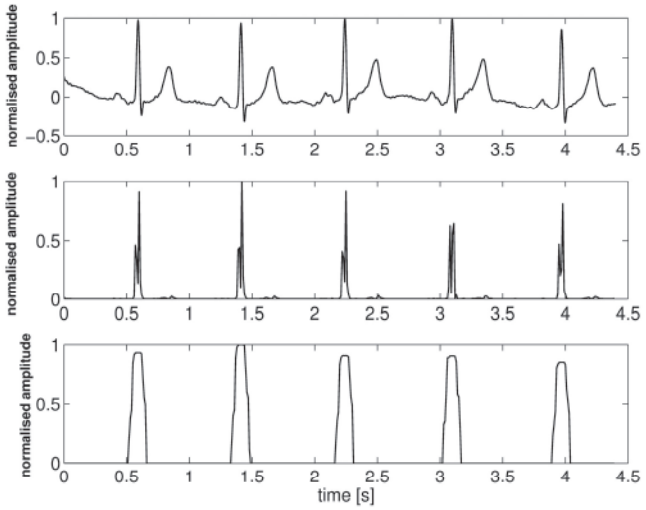

Figure 11. *Original ECG signal, ECG signal after first-derivative operator and ECG signal after application of MA filter.*

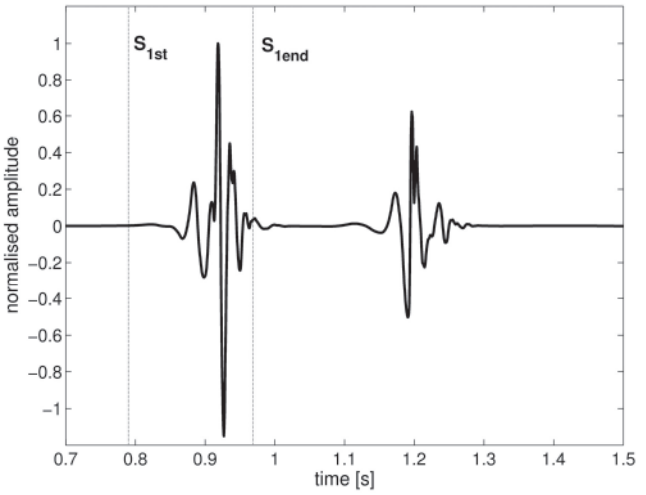

Figure 12. *Heart sound cycle showing start and end of S_1.*

implemented. This procedure is based on a first-order difference operator and a moving-average (MA) filter. The signal after the applying the first-order difference operator is (with *x(n)* being the ECG signal):

$$y(n)=\sum_{i=1}^{N}\left|x(n-i+1)-x(n-1)\right|^2(N-i+1) \tag{1}$$

after applying the MA filter:

$$f(n)=\frac{1}{M}\sum_{j=0}^{M-1}y(n-j) \tag{2}$$

where $n = 1, 2, 3,..., L$, and L is the length of the ECG signal. The original ECG signal was sampled at 2000 Hz but was digitally resampled to 100 Hz in order to speed up data processing. To prevent excessive smoothing, the window widths were set to $N = 1$ and $M = 5$. The original ECG signal, the signal after the first-derivative operator and after the MA filter application are shown in Figure 11, from top to bottom respectively. The bottom 10% of the MA-filtered signal was set to zero to remove small artifacts.

To overcome the variability in heart rates, a specific cycle's length was compared with the cycle before and after it. If a cycle's length was between pre-defined intervals, the cycle was extracted. This criterion can be defined as:

$$0.85 \le Interval \le 1.15 \tag{3}$$

where *Interval* is the ratio between the length of consecutive cycles. If a cycle adhered to the pre-defined criterion, it was extracted from the start of one QRS-peak to the start of the following QRS-peak.

Unfortunately, there is no specific event in the ECG that relates to the end of S_1, but the start of the T-wave corresponds to ventricular repolarization, which in turn results in the production of S_2. Thus the end of S_1 has to occur prior to the start of the T-wave. The ST-segment in the ECG is defined as the time between the end of the QRS-wave and the start of the T-wave. It was experimentally determined that the end of S_1 can be calculated with:

$$S_{1\,\mathrm{end}} = QRS_{\mathrm{start}} + 1.5T_{\text{S-T segment}} \tag{4}$$

where

$$T_{\text{S-T segment}} = 0.34T_{\text{R-R}}^{0.5} - 0.17T_{\text{R-R}} - 0.10 \tag{5}$$

and $T_{R\text{-}R}$ is the R-R interval duration of the ECG wave, i.e. the heart rate[17]. Figure 12 shows a phonocardiogram (one heart cycle) with the start and end of S_1 indicated as determined by this technique.

The onset of S_2 was taken as the end of the T-wave in the ECG, since this is when the ventricles start to relax and the pressure in the ventricles drop, causing the aortic and pulmonary valves to shut[8]. In calculating the end of the T-wave, the start of the QRS-complex was identified and the corrected QT interval, QT_c, was added. QT_c is equal to the QT interval divided by the square root of the R-R interval, according to Bazett's formula[19]:

$$QT_c = \frac{QT}{\sqrt{R-R}} \tag{6}$$

Figure 13 shows the extracted portion of the recorded signal with the start of S_2 indicated. As can be seen in Figure 13, S_2 as well as a diastolic murmur are present in the extracted portion of the signal (it must be cautioned not to falsely identify heart murmurs as S_2).

Finally, the end of the second heart sound cannot be attributed to any specific event in the ECG, so to determine the end of S_2 an energy envelope was calculated of the extracted signal to determine where the majority of the signal energy is situated, since this would most likely correspond with S_2. The Shannon energy was used as it intensifies the medium-intensity signals and attenuates the effect of low-intensity signals[7]. The Shannon energy envelope of the signal is shown in Figure 14. The bottom 25% of the envelope values was discarded to eliminate

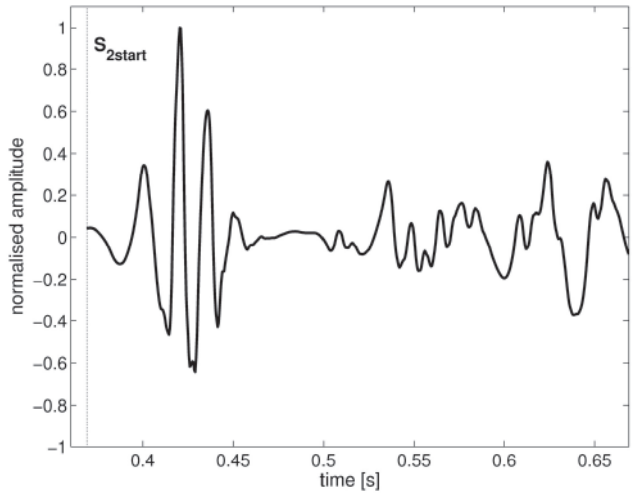

Figure 13. *Abnormal heart cycle showing start of S_2.*

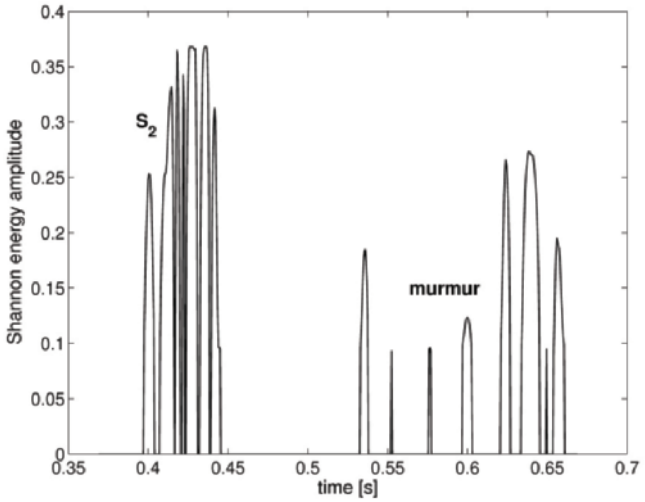

Figure 14. *Shannon energy envelope showing groupings of second heart sound.*

small noise that might interfere in the extraction process. The formula used to calculate the Shannon energy envelope was[7]:

$$E = -x^2 \log x^2 \tag{7}$$

where E is the energy envelope and x is the signal under analysis.

Peaks in the Shannon energy envelope were identified and grouped together. Two peaks were assumed to belong to the same group if the following criterion was adhered to:

$$P(i+1) - P(i) \leq 40 \ ms \tag{8}$$

where *P(i)* is an identified peak and *P(i+1)* is the next peak.

The energy of the different groups was then calculated and the group with the highest energy was extracted as S_2. This assumption proved to be sufficient in extracting the correct group as S_2. With the segmentation complete and the sounds identified, the systolic and diastolic parts of the sounds can found accurately (Figure 3).

Feature extraction

From each of the stethoscope channels (recording locations depicted in Figure 6), 70 signal features were extracted. These features were selected based on:

- discussions with cardiologists experienced in cardiac auscultation;
- previous experience in feature extraction with heart sounds;
- literature surveys; and
- trial and error experimentation with a variety of possible signal features that can relate to pathological heart sounds.

This yielded 4 x 70 = 280 features per recording instance. Features were extracted from the time and frequency domains. Time domain features included the ratio of the power between S_1 and S_2 in a specific cycle, since S_1 should be louder than S_2 at the apex of the heart and S_2 louder than S_1 at the base. To obtain the power ratio that was used as a feature, the respective power values of the extracted heart sounds were simply divided to yield the desired ratio.

To determine whether the intensity of S_1 varies from beat to beat, the power of S_1 as well as the P-R interval was compared from cycle to cycle. The calculated powers of the extracted S_1's were divided by one another to yield signal features. The P-R interval was calculated from the formulas proposed by Burke and Nasor[18]:

$$T_{\text{P-R interval}} = 0.30T_{\text{R-R}} - 0.12T_{\text{R-R}} - 0.02 \tag{9}$$

S_1 is normally longer in duration than S_2 and any deviation from this could indicate pathology[20]. Therefore, the durations of S_1 and S_2 were extracted as features.

The timing between the different components of S_2 is of great importance. During normal operation of the heart, the time gap between A_2 and P_2 increases during inspiration. This is known as splitting of the second heart sound. To detect the split of S_2, the CWT of the extracted second heart sounds were calculated. The graph that is yielded by the CWT is shown in Figure 15, with A_2 and P_2 indicated.

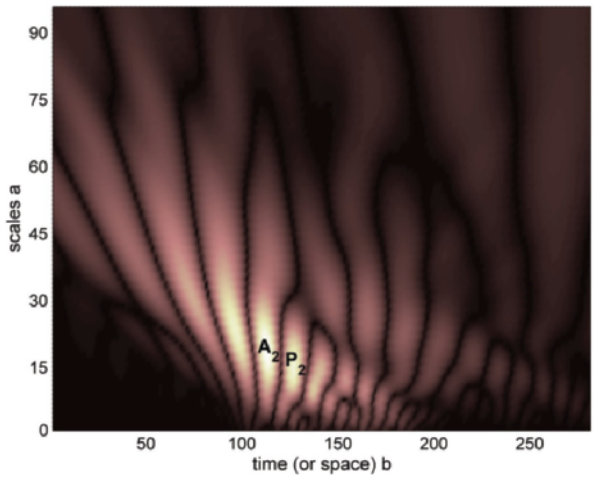

Figure 15. *CWT of S_2 with A_2 and P_2 indicated as the two highest peaks.*

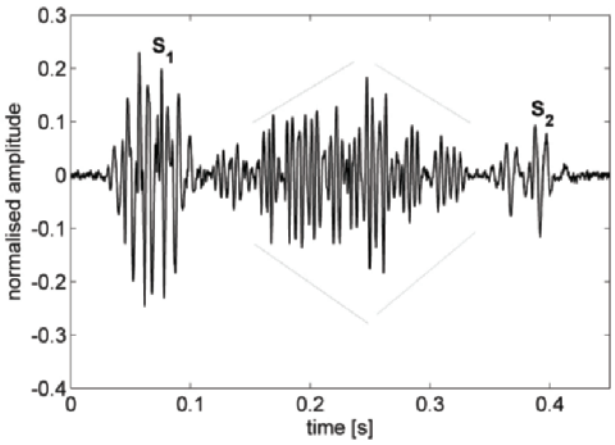

Figure 16. *Heart cycle with aortic stenosis, showing cresendo / decresendo nature of murmur.*

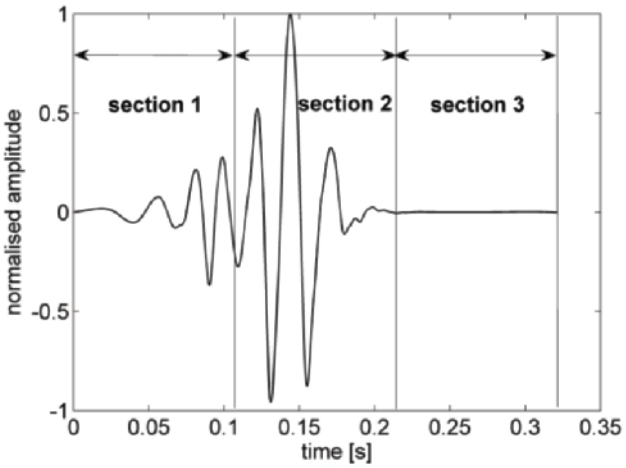

Figure 17. *Three sections of systole.*

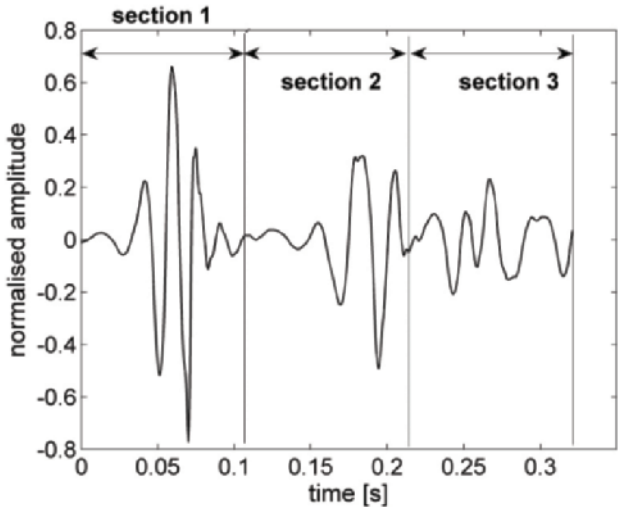

Figure 18:. Three sections of diastole.

It was decided to identify the two components of the CWT as was done by Debbal et al.[21] but to add a degree of automaticity. The db7 wavelet was used in CWT decomposition with scales from 5 to 100. It was assumed that the two highest peaks corresponded to A_2 and P_2, similar to Debbal et al.[21]. To obtain the time difference between the two components, the highest peak was identified first. It was then stepped through the entire data set to identify the second highest peak. The maximum points were identified and subsequently set to zero until two maxima differed by 10 ms or more. This maximum was then identified as the second peak. The absolute value of the time difference between the two components was then taken as the time difference between A_2 and P_2.

The shape of the systolic and diastolic murmurs is indicative of the type of murmur that is present in the heartbeat of a patient, for instance as depicted in Figure 16 for aortic stenosis. To calculate the shape of the murmur, systole and diastole were split into three sections each (Figure 17 and Figure 18), and the Root Mean Square (RMS) value of each section was calculated. Depending on whether the values increased, decreased or stayed more or less the same from section to section, the shape of the murmur can either be described as crescendo, decrescendo, crescendo-decrescendo or plateau. The maximum frequency of each of the extracted sections of systole and diastole were also considered as a feature. The maximum frequency corresponds to the frequency in the FFT that has the maximum amplitude.

The presence of extra heart sounds such as the ejection sound, midsystolic click and opening snap was also included as signal features. To check whether any of these extra sounds were present, it was argued that within the interval in which the sounds would occur, the power of the sections would be higher if these sounds were present than if they were absent. The systolic and diastolic regions of each cardiac cycle were identified and broken up into different sections, as shown in Figure 19, to search for these extra sounds. ES refers to the ejection sound, MC to the midsystolic click and OS to the opening snap.

For frequency domain features, an FFT frequency ratio was calculated for S_1 and S_2. The magnitudes of the Fourier coefficients in the frequency band from 0 Hz to 100 Hz were summed, as well as the magnitudes of the Fourier coefficients in the frequency band from 100 Hz to 800 Hz and divided by one another to yield a frequency ratio.

Feature reduction and ANN

The procedures described above yielded a large number of signal features, and had to be reduced for the ANN development. The input feature space of ANNs is often reduced to simplify the ANN structure, reduce training time and improve the classification capabilities. The Statistical Overlap Factor (SOF) was used to identify the features that exhibited the greatest degree of separation between the recordings of normal and abnormal heart sounds. The SOF is defined as[22]:

$$SOF = \left| \frac{\bar{x}_1 - \bar{x}_2}{(\sigma_1 + \sigma_2)/2} \right| \tag{10}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the means of distributions x_1 and x_2, and σ_1 and σ_2 are the respective standard deviations. The higher the SOF, the better the degree of separation between

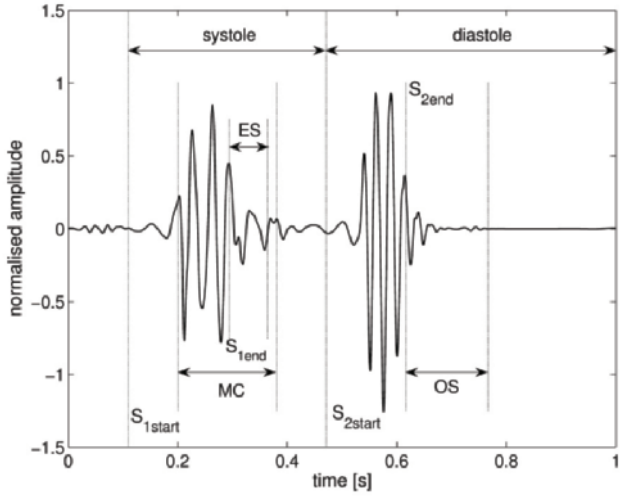

Figure 19. *Normal heart cycle showing search areas for extra heart sounds.*

the two distributions. It was determined through trial and error simulations with large and small input spaces, that the following six features identified by the SOF yielded the best classification results with the ANN:

$$AR = \frac{\text{amplitude gradient of first half of diastole}}{\text{amplitude gradient of second half of diastole}}$$

$$AS = \frac{\text{amplitude gradient of first half of systole}}{\text{amplitude gradient of second half of systole}}$$

$$MR = \frac{\text{RMS(systole)}}{\text{RMS(diastole)}}$$

$$CF = \text{CREST FACTOR(entire heart sound cycle)}$$

$$MR = \frac{\text{RMS(last third of systole)}}{\text{RMS(last third of diastole)}}$$

$$MS = \frac{\text{RMS(entire systole)}}{\text{RMS(last third of diastole)}}$$

Best results were achieved when these six features were used for all four stethoscope locations, hence yielding an input space of 24 features. An ANN was hence constructed that consisted of 24 inputs, 30 hidden neurons and one output neuron. The network was constructed to give an output of 0 for a healthy heart and a 1 for a pathological heart.

All nodes used the logarithmic tangent as the activation function, except for the output node, which implemented a linear activation function. The network was trained with the backpropagation algorithm and used an adaptive learning rate. The cost function used was the mean-square-error function with regularization, with the regularization parameter set to 0.5.

Because the number of recordings were limited, cross-validation was used to validate and test the ANN. For each training instance, the recordings from one person is excluded, and then used as a test sample after training. The training instances were repeated 500 times with a new, random network initialization for each recorded person, and the average network output was used to calculate the network's overall performance.

The overall result was a sensitivity of 84%, a specificity of 86% and a total network accuracy of 85%.

Discussion

The stethoscope recordings were quite noisy and it was found that in some instances the denoising procedure did not produce good results. If the wavelet threshold was set too high, some necessary information was discarded, such as S_1, S_2 or murmur information. Low- and high-pass filters could not be used to correct this, since the noise frequency spectrum coincided with the frequency spectrum of the heart sounds. The analysis of the data was completely automated, (i.e. it did not require the adjustment of analysis parameters for individual recordings). Hence all the data had to be denoised in the same manner. The wavelet denoising method uses a threshold for the denoising procedure, and it was determined experimentally that 0.3×SURE provided the best results for most of the recorded heart sounds.

The reason for the poor recordings could be that in some instances the stethoscopes did not make sufficient contact with the patient's body and this resulted in motion artifacts being recorded as the patient breathes. The auscultation jacket was constructed based on anthropometric data and proved to fit most of the patients that had normal heart sounds, but in some instances did not fit well on patients that suffered from heart disease. The reason for that being that heart disease patients are sometimes of a smaller build and could suffer from chest deformities such as pectus excavatum. The aim of this study was in part to demonstrate the potential of the auscultation jacket, and it is obvious that the device will have to be constructed in various different sizes in order for the concept to be completely feasible.

Another reason for recording excessive noise might be that the stethoscopes were not placed on the exact auscultation positions. This could also be contributed to the design of the jacket. The heart sound might then still be present but is deeply buried in background noise so that too much denoising simply removes the recorded heart sound. Research and development is currently underway to improve the design of the jacket, and manufacture it in various sizes. The long-term vision is that the jacket could become a tool for screening patients for heart pathologies in rural healthcare, through telemedicine.

The procedure that extracted the time duration between the two components of S_2 is not ideal. More constraints need to be identified to make the identification easier. The breathing cycle also needs to be incorporated in the constraints since the gap between A_2 and P_2 is very dependent on the breathing cycle.

The ECG that was recorded simultaneously with the heart sound data produced artifacts in some instances. It is

speculated by the authors that these artifacts may have been produced by the customized ECG system that built into the auscultation jacket. The ECG used in this study was built in-house simply to produce the QRS-complex so that the start of S_1 could be identified, and hence it is not a diagnostic tool such as a commercial ECG. In instances where artifacts were present in the recorded ECG, the ECG extraction algorithm was unsuccessful. As a result, reliable information could not be extracted from these recordings. This led to the problem that the recordings from several participants had to be discarded, with the consequence was that the results presented in this paper are based on a relatively small data set. However, the proposed methods yielded good results and the authors are confident that the results can only improve by adding more data for training, and also addressing the hardware problems.

The outcome of this research was very dependent on the volunteers that participated. As explained above, many recordings were of such a poor quality that they could not be used. It is proposed that a large set of patients (100 or more), that cover a broad spectrum of auscultatory abnormalities, be identified and their heart sounds recorded with improved hardware. Subsequently, the algorithms and features presented in this study may need further refinement along with this new data set, which would hopefully include a larger variety of heart pathologies.

Conclusion

An ANN-based system has been developed that is capable of differentiating between normal and certain auscultatory abnormalities. The system uses features extracted from heart sounds recorded at multiple locations on the thorax as input to the classification system. The classifier showed a sensitivity of 84% and a specificity of 86%. Data was recorded with a low cost and easy to use auscultation jacket, and even though the sets presented in this paper are limited, it shows great potential for such a device along with diagnostic software. In its current form, the system is not ready for clinical deployment, but further research and refinement will be directed towards deploying such a system in a rural healthcare scenario where there are limited resources and many patients, which can benefit a lot from such technology.

Acknowledgements

The authors would like to thank Tygerberg Hospital, especially the TREAD Research Unit for allowing us to use their facility for the recording procedures. A special word of thanks towards Dr. H.L. Koekemoer for his support and advice, and the NRF under grant number FA2005022500002 for financial support.

References

1. International Cardiovascular Disease Statistics [Online], Available at: http://www.americanheart.org/downloadable/heart/1140811583642InternationalCVD.pdf, [Accessed 24 April 2006], 2006.
2. Bradshaw, D., Groenewald, P., Laubscher, R., Nannan, N., Nojilana, B., Norman, R., Pieterse, D., Schneider, M., Bourne, D., Timaeus, I., Dorrington, R., and Johnson, I., *Initial burden of disease estimates for South Africa, 2000*, South African Medical Journal, 93, 682-688, 2003.
3. Leeder, S., Raymond, S. and Greenberg, H., A race against time: *The challenge of cardiovascular disease in developing economies* [Online], Available at: http://www.ahpi.health.usyd.edu.au/pdfs/colloquia2004/leederracepaper.pdf, [Accessed 25 April 2006], 2004.
4. Essop, M.R., Nkomo, V.T. *Rheumatic and non-rheumatic valvular heart disease: epidemiology management and prevention in Africa*, Circulation, 112, 3584-3591, 2005.
5. Obaidat, M., *Phonocardiogram signal analysis: techniques and performance comparison*, Journal of Medical Engineering & Technology, 17, 221-227, 1993.
6. March, S., Bedynek, J. and Chizner, M., *Teaching cardiac auscultation: effectiveness of a patient-centered teaching conference on improving cardiac auscultatory skills*, Mayo Clinical Proceedings, 80, 1443-1448, 2005.
7. Liang, H., Lukkarinen, S. and Hartimo, I., *Heart sound segmentation algorithm based on heart sound envelogram*, in: IEEE, ed., Computers in Cardiology (IEEE, Lund, 1997) 105-108, 1997.
8. Rangayyan, R.M., *Biomedical Signal Analysis: A Case-Study Approach*, New York, Wiley-Inter-Science, 2002.
9. Andrisevic, N., Ejaz, K., Rios-Gutierrez, F., Alba-Flores, R., Nordehn, G. and Burns, S., *Detection of heart murmurs using wavelet analysis and artificial neural networks*, Journal of Biomechanical Engineering, 127, 899-904, 2005.
10. Cathers, I., *Neural network assisted cardiac auscultation*, Artificial Intelligence in Medicine, 7, 53-66, 1995.
11. DeGroff, C.G., Bhatikar, S., Hertzberg, J., Shandas, R., Valdes-Cruz, L. and Mahajan, R.L., *Artificial neural network-based method of screening heart murmurs in children*, Circulation, 103, 2711-2716, 2001.
12. de Vos, J.P., Automated Pediatric Cardiac Auscultation, Master's thesis, Electric & Electronic Engineering, Stellenbosch University, Stellenbosch, South Africa, 2005.
13. Gupta, C.N., Palaniappan, R., Swaminathan, S. and Krishnan, S.M., *Neural network classification of homomorphic segmented heart sounds*, Applied Soft Computing, 7, 286-297, 2007.
14. Koekemoer, H.L. and Scheffer, C., *Heart Sound and Electrocardiogram Recording Devices for Telemedicine Environments*, Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2008), p. 4867 – 4870, Vancouver, Canada, 20-25 August 2008,.
15. Bentley, P.M., Grant, P.M. and McDonnell, J.T.E., *Time-frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds*, IEEE Transactions on Biomedical Engineering, 45, 125-128, 1998.
16. Ölmez, T. and Dokur, Z., *Classification of heart sounds using an artificial neural network*, Pattern Recognition Letters, 24, 617-629, 2003.
17. Messer, S.R., Agzarian, J. and Abbott, D. *Optimal wavelet denoising for phonocardiograms*, Microelectronics Journal, 32, 931-941, 2001.
18. Walker, J.S., *A Primer on Wavelets and their Scientific Applications*, New York, Chapman & Hall/CRC, Chapter 2, 58-65, 1999.
19. Burke, M.J. and Nasor, M., *The time relationships of the constituent components of the human electrocardiogram*,

Journal of Medical Engineering & Technology, 26, 1-6, 2002.
20. Bazett, H.C., *An analysis of the time-relations of electrocardiograms*, Heart, 7, 353-370, 1920.
21. Werner, L., Pitts, B. and Gilsdorf, D., *Heart sounds: An interactive auscultation program,* CD-ROM to accompany Bates' Guide to Physical Examination, Philadelphia, Lippincott Williams & Wilkins, 2005.
22. Debbal, S. M. and Bereksi-Reguig, F., *Heartbeat sound analysis with the wavelet transform*, Journal of Mechanics in Medicine and Biology, 4, 133-141, 2004.
23. Mdlazi, L., Marwala, T., Stander, C., Scheffer, C. and Heyns, P., *The principal component analysis and automatic relevance determination for fault identification in structures*, in: Tom Proulx, ed., Proceedings of the 21st International Modal Analysis Conference (IMAC 21) (Society for Experimental Mechanics, Inc., Florida) 37-42, 2003.