Computers in Biology and Medicine 38 (2008) 221–233

Computers in Biology and Medicine

www.intl.elsevierhealth.com/journals/cobm

A computer-aided MFCC-based HMM system for automatic auscultation

Sunita Chauhan[a,*], Ping Wang[a], Chu Sing Lim[a], V. Anantharaman[b]

[a]School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 - Nanyang Avenue, Singapore 639798, Singapore
[b]Department of Emergency, Singapore General Hospital, Singapore

Received 25 May 2006; accepted 12 October 2007

Abstract

Auscultation, the act of listening to the sounds of internal organs, is a valuable medical diagnostic tool. Auscultation methods provide the information about a vast variety of internal body sounds originated by various organs such as heart, lungs, bowel, vascular disorders, etc. In this study, a cardiac sound registration system has been designed incorporating functions such as heart signals segmentation, classification and characterization for automated identification and ease of interpretation by the users. Considering a synergy with the domain of speech analysis, the authors introduced Mel-frequency cepstral coefficient (MFCC) to extract representative features and develop hidden Markov model (HMM) for signal classification. This system was applied to 1381 data sets of real and simulated, normal and abnormal domains. Classification rates for normal and abnormal heart sounds were found to be 95.7% for continuous murmurs, 96.25% for systolic murmurs and 90% for diastolic murmurs by a probabilistic comparison approach. This implies a high potential for the system as a diagnostic aid for primary health-care sectors.

Keywords: Heart sounds; Mel-frequency cepstral coefficient (MFCC); Hidden Markov model (HMM); Phono-cardiogram

1. Introduction

Auscultation is a valuable means for rendering vast amount of information concerning various sounds originated by internal body organs such as heart, lungs, bowel and vascular disorders. For centuries, cardiovascular diseases remain one of the major health concerns. Early detection of heart diseases and accurate diagnosis of related conditions comprise a significant medical research area. There are several diagnostic methods for analyzing the functioning of the heart, ranging from auscultation, electro-cardiogram (ECG), echo-cardiogram, ultrasound, etc. Auscultation is the most common and cost-effective non-invasive technique and heart sound analyses can provide important information relating to heart condition for diagnosis of cardiovascular disorders [1–3]. Phonocardiogram (PCG) can provide valuable information concerning the integrity and function of the heart valves and hemodynamic mechanisms. Stethoscope is a basic device for hearing internal sounds and has become a part of the identity of medical staff. The information acquired by a traditional stethoscope is, however, relative and qualitative in nature since the differentiation of signals picked-up by the sensor during manual interpretation is limited by human perception and varies with personal aptitude and training. This may result in inaccurate or insufficient information due to the inability of the user to discern certain complex, low-level, short duration or rarely encountered abnormal sounds. It is, thus, desirable to enhance the diagnostic ability by electronically processing the stethoscope signal output at the sensor level and providing a visual display to the physician for a better comparative study.

During the last decade, efforts have been made in this direction and electronic stethoscopes are now available commercially with phono-cardiograph display. The usage of such devices, however, is less common due to the involvement of cumbersome instrumentation and complex/additional skills. With a rapid advancement of electronic instrumentation and digital processing techniques as well as an improved understanding of the genesis of heart sounds, many characteristics and critical features of heart sounds can be analyzed very accurately.

* Corresponding author. Tel.: +65 6 790 5520; fax: +65 6 792 4062.
E-mail address: mcsunita@ntu.edu.sg (S. Chauhan).

0010-4825/$ - see front matter
doi:10.1016/j.compbiomed.2007.10.006

In time domain, the first heart sound ($S1$) period is ~ 110 ms from the ECG R wave peak time. The period from 40 ms before to 60 ms after the peak time can be considered as the second heart sound, $S2$ [4]. The period between the end of the first heart sound and beginning of the second heart sound is determined as the systolic period. Liang et al. [5,6] identified $S1$ and $S2$ based on the comparative intervals of diastole and systole periods. Several research efforts have been made to analyze heart sounds in frequency domain as well by using fast Fourier transform (FFT). By studying the specific frequencies of heart sounds, it is found that the first and second heart sounds fall, respectively, in the frequency range of 80–120 Hz and 120–150 Hz. The third and fourth heart sounds are found to be of low frequencies in the range of 70–90 Hz and 50–70 Hz, respectively [7]. Average murmurs are found to have usually higher frequencies in the range of 100–600 Hz [1,8]. Time–frequency representative algorithms have been applied in heart sound analysis such as short-time Fourier transform (FT) [9,10] and wavelet transform [11,12], due to the non-stationary nature of the signals. Pattern recognition algorithms, such as using neural networks, have been developed in heart sound classification [13–16]. Sava et al. [13] presented the application of a single layer perceptron for classification of normal and malfunctioning bio-prosthetic heart valves. Multilayer feed-forward network for classifying normal and abnormal heart sounds [15,16] and multilayer feed-forward network trained by back propagation have been used to distinguish between first and second heart sounds [14]. However, in the latter method, a lack of quantitative and systematic implementation for heart sound analysis was observed. This could be due to a statistical requirement of a large number of training data by neural network to achieve good performance.

In the present study, the authors have introduced an automated heart sound classification system based on Mel-frequency cepstral coefficient (MFCC) and hidden Markov model (HMM). The system works in several systematic steps preceded by a pre-processing procedure including data acquisition, filtering, normalization and segmentation. The paper is organized as follows: an introduction in Section 1 is followed by an overview of the experimental methodology and a brief description of multi-stage processing such as data acquisition pre-processing, feature extraction, feature classification and characterization in Section 2 (further details can be found in [17]. Results and discussions are presented in Section 3 followed by conclusions of the study in Section 4.

2. Methodology

At NTU, we have developed an automated stethoscope system with the ability to freeze and store the heart sound signals for quantitative and off-line consultation that would aid the physicians to build a consensus view, leading to more contemplative examination. As shown in the schematic diagram of Fig. 1, the system consists of five sequential steps: (1) data acquisition, (2) signal pre-processing, (3) feature extraction, (4) model training and recognition and (5) characterization. The functions of these steps and procedures are described in the following subsections:

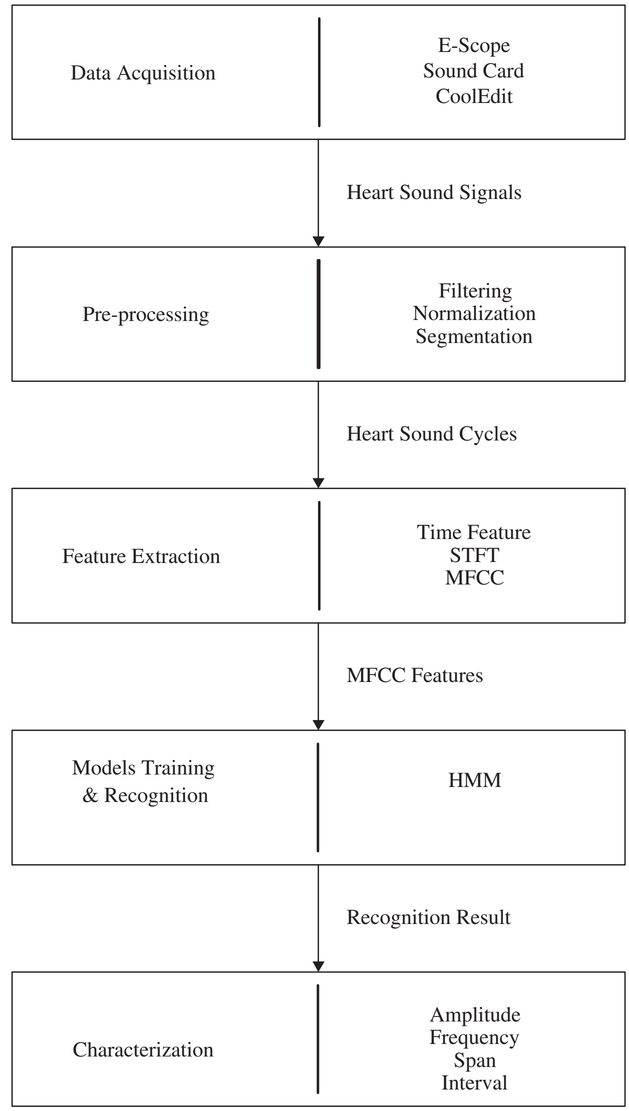


Fig. 1. Schematic diagram of the system.

2.1. Data acquisition

During the data acquisition stage heart sound signals are directly collected from patients and saved as waveforms. The procedure comprises five functional blocks as shown in Fig. 2. For experimental analysis in this study, 41 samples were analyzed: the sample group consisted of 20 normal subjects, six patients with reported heart conditions with continuous murmurs, four patients with diastolic murmurs and 11 with systolic murmurs. The heart sound signals were acquired by the Cardionics E-Scope® Electronic Stethoscope (p/n: 718–7120: Cardionics Inc., USA), whose output was connected to a sound card with 16-bit A/D conversion resolution. Heart sound signals were sampled at 8 kHz using CoolEdit 2000 (Syntrillium Software Corporation, USA). Fig. 3 shows specific/preferred positions (P1, P2, P3 and P4) for data acquisition using the E-scope and sample waveforms using CoolEdit.

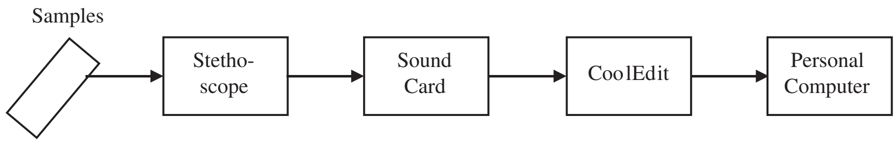


Fig. 2. Block diagram of data acquisition.

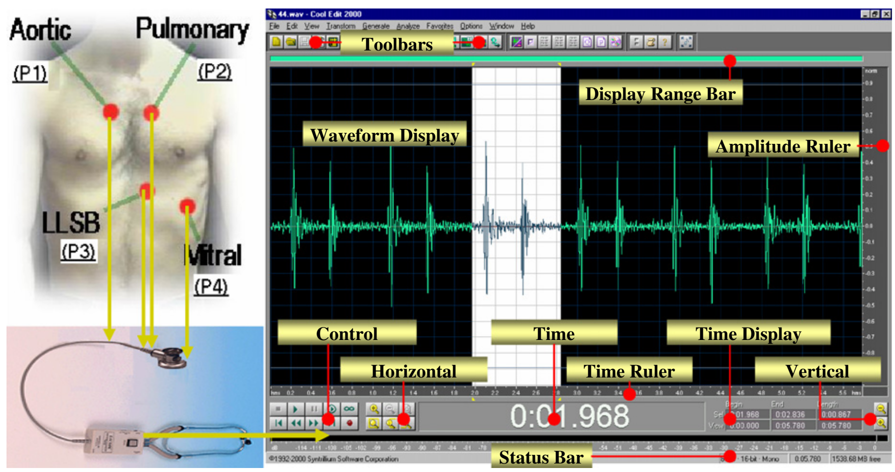


Fig. 3. Block diagram showing data acquisition steps using E-scope and Cooledit (P1, aortic area; P2, pulmonary area; P3, tricuspid area (lower left sternal border); P4, mitral area).

2.2. Signal pre-processing

In this stage, the heart sound signals are prepared for smooth playback, followed by noise filtration and segmentation. During signal pre-processing, the following steps were included:

Filtering: Heart sound signals were filtered using a low-pass filter to remove unwanted high-frequency components followed by a digital finite impulse response (FIR) filter with a symmetric Hamming window to reduce the undesirable ripples and ringing. Median filtering was applied to smooth and sharpen the boundaries of the signals while removing noises.

Normalization: In data acquisition, different sampling and acquisition locations normally result in a signal variation. Thus, the heart sound signals were normalized using the following equation [5,6]:

$$\mathrm{HS}_{\mathrm{norm}}(n) = \frac{\mathrm{HS}(n)}{\max(|\mathrm{HS}(n)|)}, \tag{1}$$

where $\mathrm{HS}(n)$ is the raw heart sound signals and $\mathrm{HS}_{\mathrm{norm}}(n)$ indicates the normalized signals. It tended to transform the signals to a scale of $[-1, 1]$ so that the expected amplitude of the signal is not affected from the data acquisition locations and different samples.

Segmentation: After getting the filtered normalized signals, heart sounds were segmented into cycles using the procedure shown in Figs. 4 and 5. Shannon energy (Eq. (2)) was used to identify $S1$ and $S2$ [5,6] by performing a threshold function on the envelope of the heart signals:

$$\mathrm{SE} = -\mathrm{HS}^2_{\mathrm{norm}}(n)\,\log\,\mathrm{HS}^2_{\mathrm{norm}}(n). \tag{2}$$

In all the cycles, the two sounds, $S1$ and $S2$ can be clearly distinguished by a small pause [18]. The two adjacent peaks of $S1$ and $S2$ are about 50 ms apart, thus a limit window of 50 ms was used to scan throughout the whole set of Shannon energy signals. Due to the complexity of constituent signals in factual heart sound cycles, and in order to avoid any suspected data loss, two additional procedures were included based on time limit, namely rejecting and recovering procedures. When more than one peak appeared in a time-limit window which meant an extra peak was selected, their energies were compared and the one with less energy was rejected. With no peak in a time-limit window, the threshold was reduced until the lost peak was recovered. Based on pause lengths, $S1$ and $S2$ were identified separately. Thus, raw signals were segmented into cycles. The segmentation procedure of heart sound record with $S4$ is shown in Fig. 5. For further details on the procedural steps during segmentation, please refer to [17].

2.3. Feature extraction

In all recognition systems, signal processing is carried out to convert the raw signals to some type of parametric representation. This parametric representation is then used for further analysis and processing and is called a feature. After collecting the data, a database is created from which required features of the data could be extracted for processing by HMM models. Three types of features were extracted and compared: (1) time-domain feature, (2) Short-time Fourier frequency transform (STFT) and (3) MFCC.

Time-domain features: These include the filtered and normalized signals in which the independent variable is time and

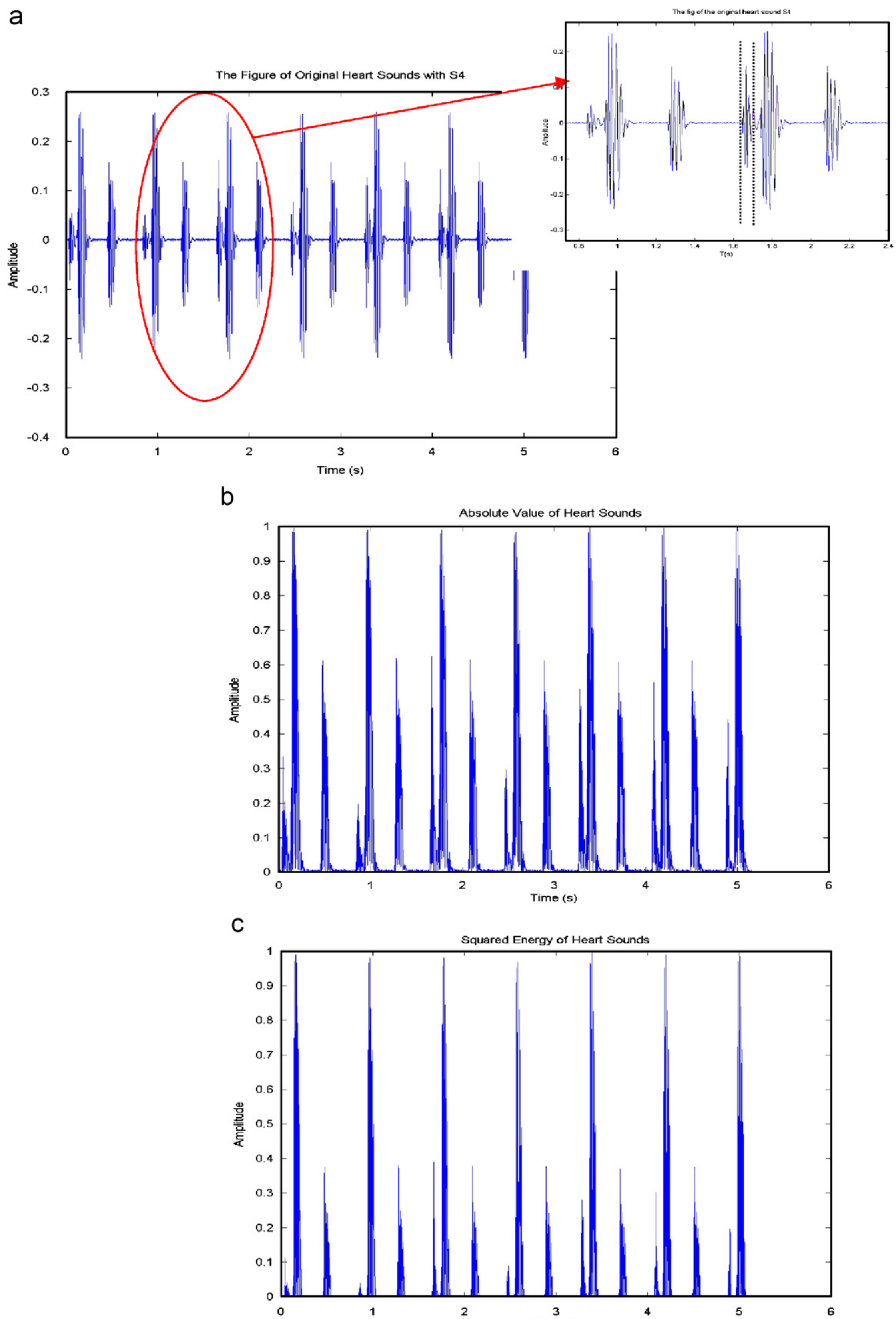


Fig. 4. (a) Raw heart sounds with $S4$; (b) absolute value of heart sounds with $S4$; (c) squared energy of heart sounds with $S4$; (d) Shannon entropy of heart sounds with $S4$; (e) Shannon energy of heart sounds with $S4$; and (f) normalized average Shannon energy of heart sounds.

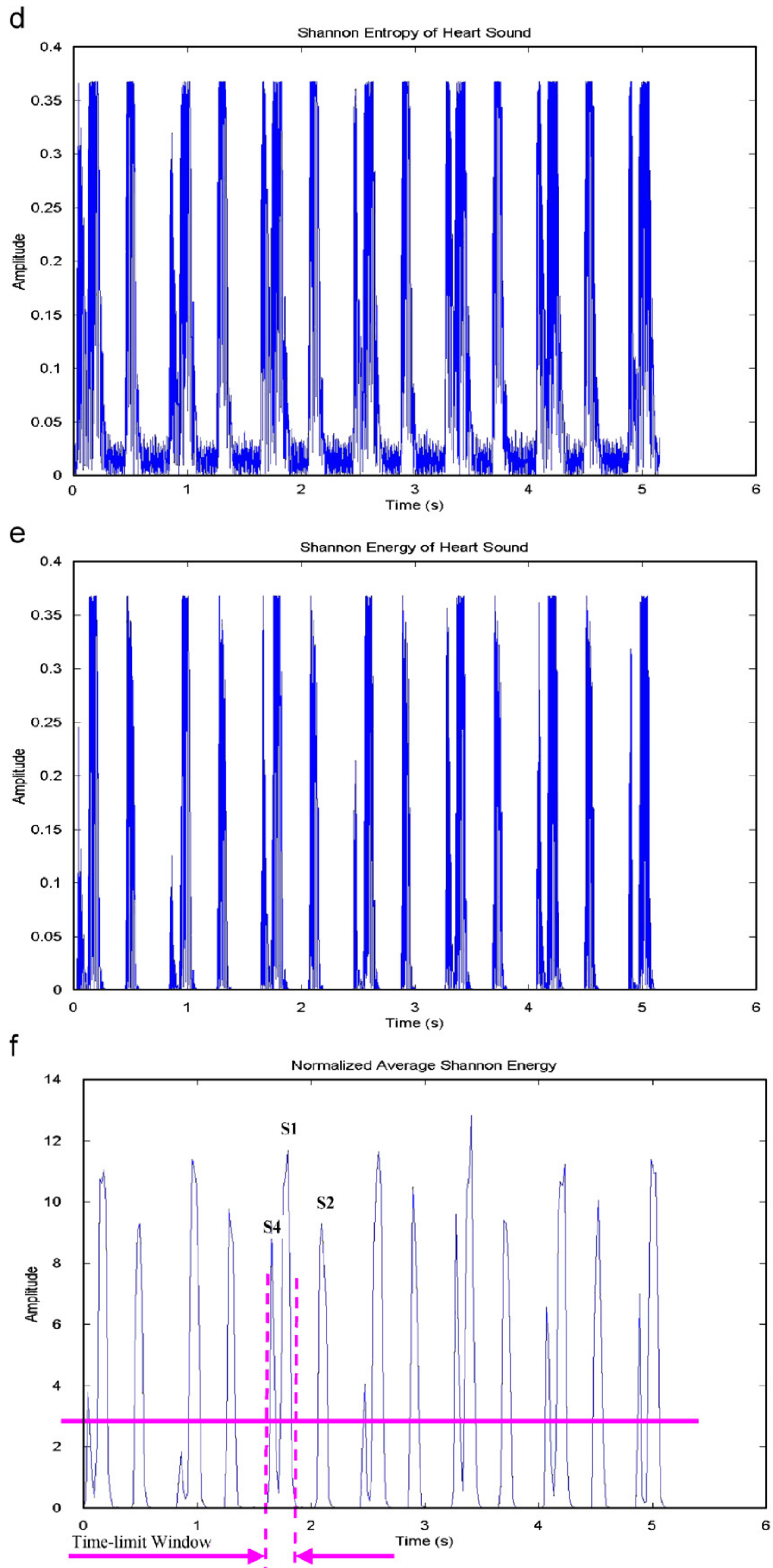


Fig. 4. (*continued*).

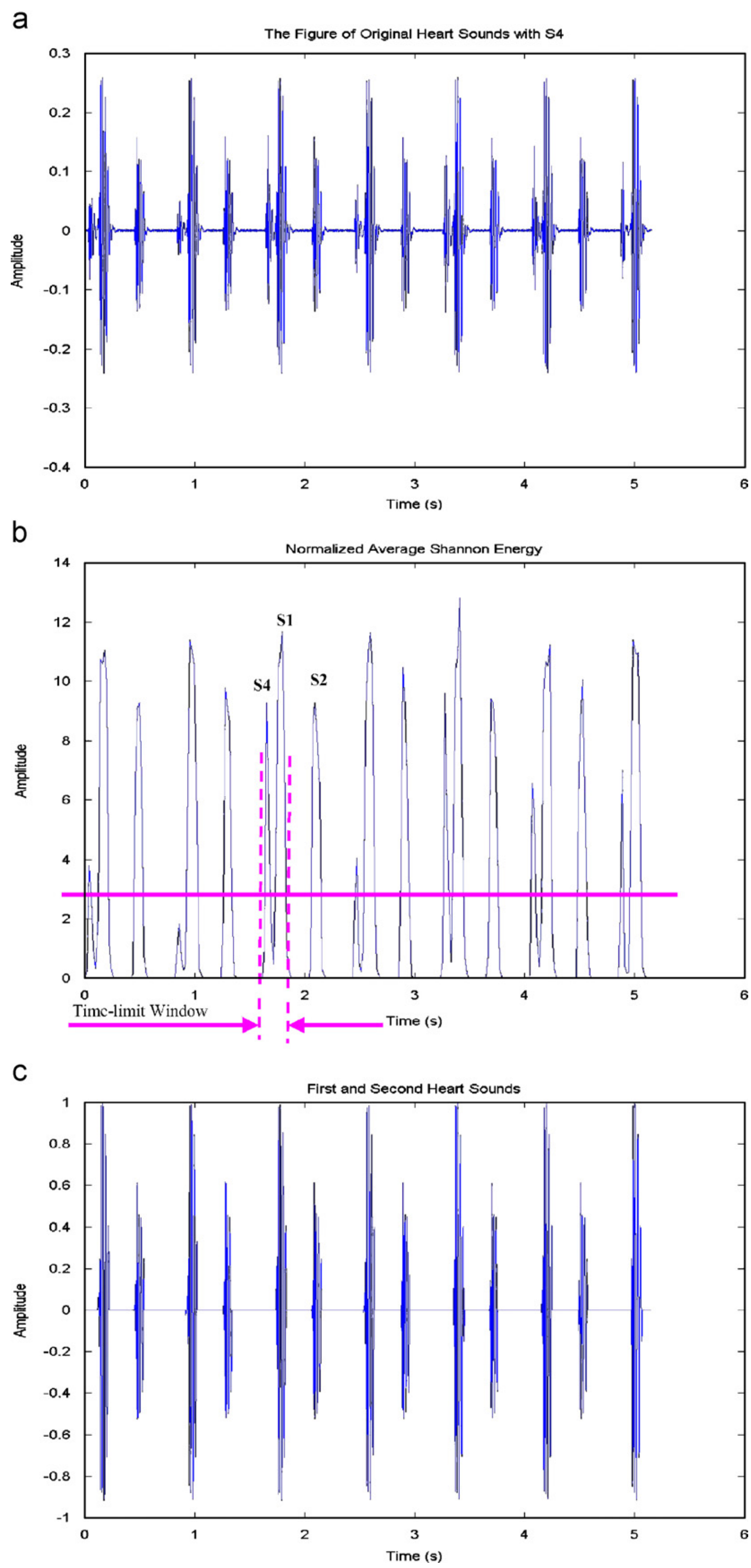
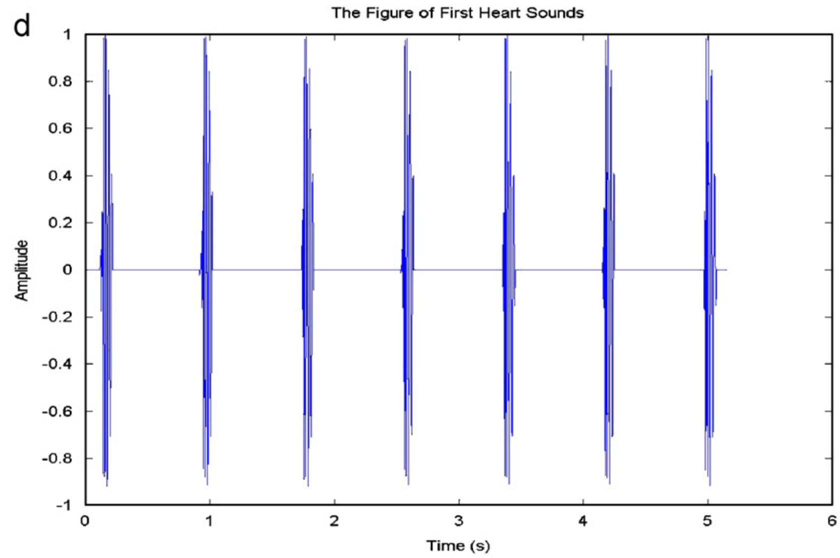
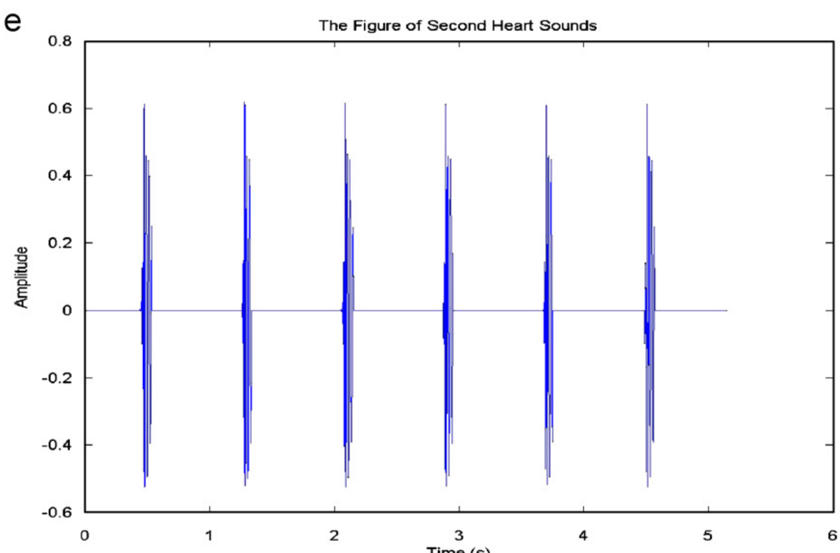


Fig. 5. (a) Raw heart sounds with $S4$; (b) Shannon energy of heart sounds with $S4$; (c) segmented $S1$ and $S2$; (d) separated $S1$; (e) separated $S2$; and (f) segmented second cycle.





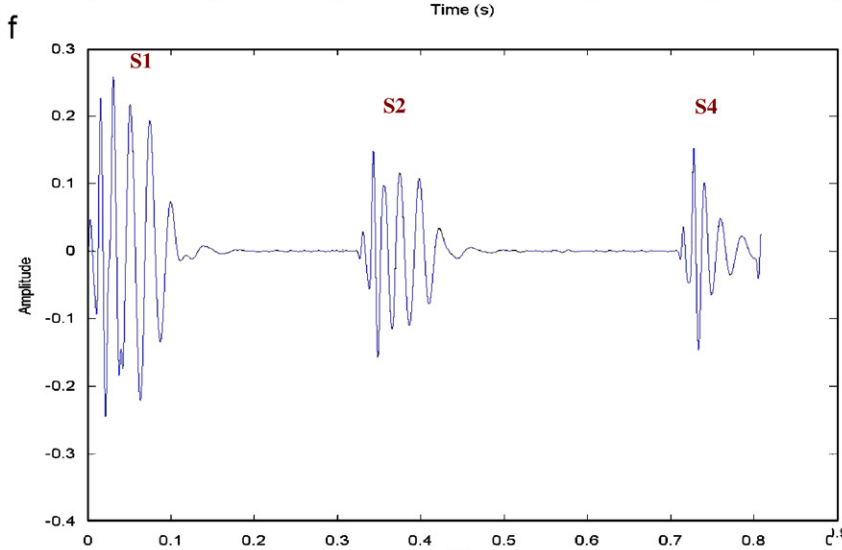


Fig. 5. (*continued*).

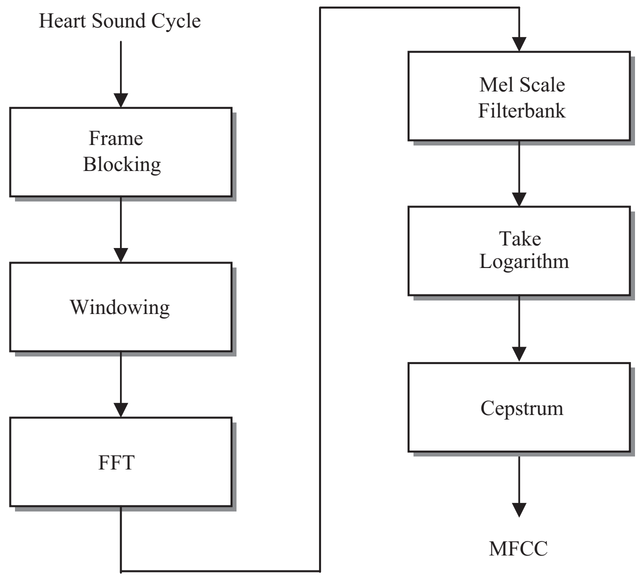


Fig. 6. Procedure of MFCC.

the dependent variable is the amplitude. Time domain feature, however, is unable to reveal the distinguished information hidden in the frequency content.

Short-time Fourier frequency transform: STFT was obtained from the usual FT by multiplying the time signal $x(t)$ by an appropriate sliding Hamming window $w(t)$:

$$X(t, \omega) = \int x(t) w(\tau - t) \mathrm{e}^{-\mathrm{j}\omega\tau} \, \mathrm{d}\tau. \tag{3}$$

Mel-frequency cepstral coefficient: MFCC coefficients were obtained by taking a discrete cosine transform (DCT) of the logarithmic spectrum scale after it was warped to the Mel scale [19] by using Eq. (4), which is similar to perceptual linear predictive analysis of sound signals [20]:

$$\mathrm{mel}(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right). \tag{4}$$

A flowchart showing the procedure carried out for calculating the MFCCs is delineated in Fig. 6. In a frame blocking step the continuous heart sound signals were blocked into frames of 25 ms with 10 ms overlapping. With 8 K sampling frequency, the frame size N was 200 and the overlapping size M was 80. After dividing the signal into frames, the next step was to assign a window to each individual frame. The Hamming window was used to minimize the spectral distortion by tapering the signal to zero at the beginning and end of each frame. FFT was then carried out on the signal window to convert each frame of N samples from the time domain into a frequency domain.

The perception of the human ear does not follow a linear scale. Thus, for a signal with an actual frequency, f, measured in Hz, a subjective pitch is measured on the 'Mel' scale. Mel scale was used to transform the power spectrum to compute a Mel-warped spectrum. In order to simplify the spectrum without significant loss of data, the Fourier transformed signal was passed through a set of band-pass filters. This filter bank had a triangular band-pass frequency response. The triangular filters used were non-uniform at the original spectrum and more uniformly distributed at the Mel-warped spectrum Each filter in the bank was multiplied by the spectrum so that only one single value of magnitude per filter was returned. This value reflected the sum of amplitudes in a particular filter band and thus reduced the precision to the level of human ear. The modified spectrum of heart sounds thus consisted of the output power of these filters. Logarithm of the Mel spectrum coefficients was then taken to compress the coefficients above 1000 Hz and also to compress the magnitude with low frequencies. In the final step, the logarithmic Mel spectrum was converted back to time domain. This resulted in the MFCC. Normal heart sounds before MFCC and the MFCC coefficients are shown in Fig. 7.

2.4. Model training and recognition

The heart sounds collected from clinical samples comprised four types: normal, continuous murmurs, diastolic murmurs and systolic murmurs. Data from the public domain involved six types of heart sounds: ejection click, opening snap, split $S1$, split $S2$, $S3$ and $S4$. These various types of heart sounds were used for HMM modeling. As such, 10 models of heart sounds were trained individually by HMM and unknown input heart sound signals were classified automatically by the trained models. A typical left-to-right HMM [21,22] was selected to represent heart sound signals since the signals were one-dimensional time varying series. In the training procedure, model parameters were estimated using maximum likelihood formulas. This approach, called Baum–Welch re-estimation [21], produces new parametric estimates, which have an equal or greater likelihood of having generated the training data, in this case, the 12 MFCC features.

An iterative procedure using the forward–backward algorithm [21,23] was used to reduce the computation time. For each iteration the forward and backward probabilities were computed and the results were used to compute new parameter estimates. The recognition correct percentage (RCP), which used the known testing data for classification, was employed to evaluate the trained models.

$$\mathrm{RCP} = \text{Accurately identified cycles/Total cycles.} \tag{5}$$

Each HMM model was trained by one reported condition of heart sounds. After training the models, probabilities were computed for each of the trained models and recognition results were obtained by choosing the most likely model. Forward–backward procedure was used for calculating the probability to reduce the calculations instead of direct computation (i.e. by enumerating all possible state sequences and summing the probability of generating the observation sequence from each of the sequences). A 5-state HMM is shown in Fig. 8.

2.5. Characterization

After recognizing the input heart signals, each component of the input data was identified (such as $S1$, $S2$, murmurs, etc.)

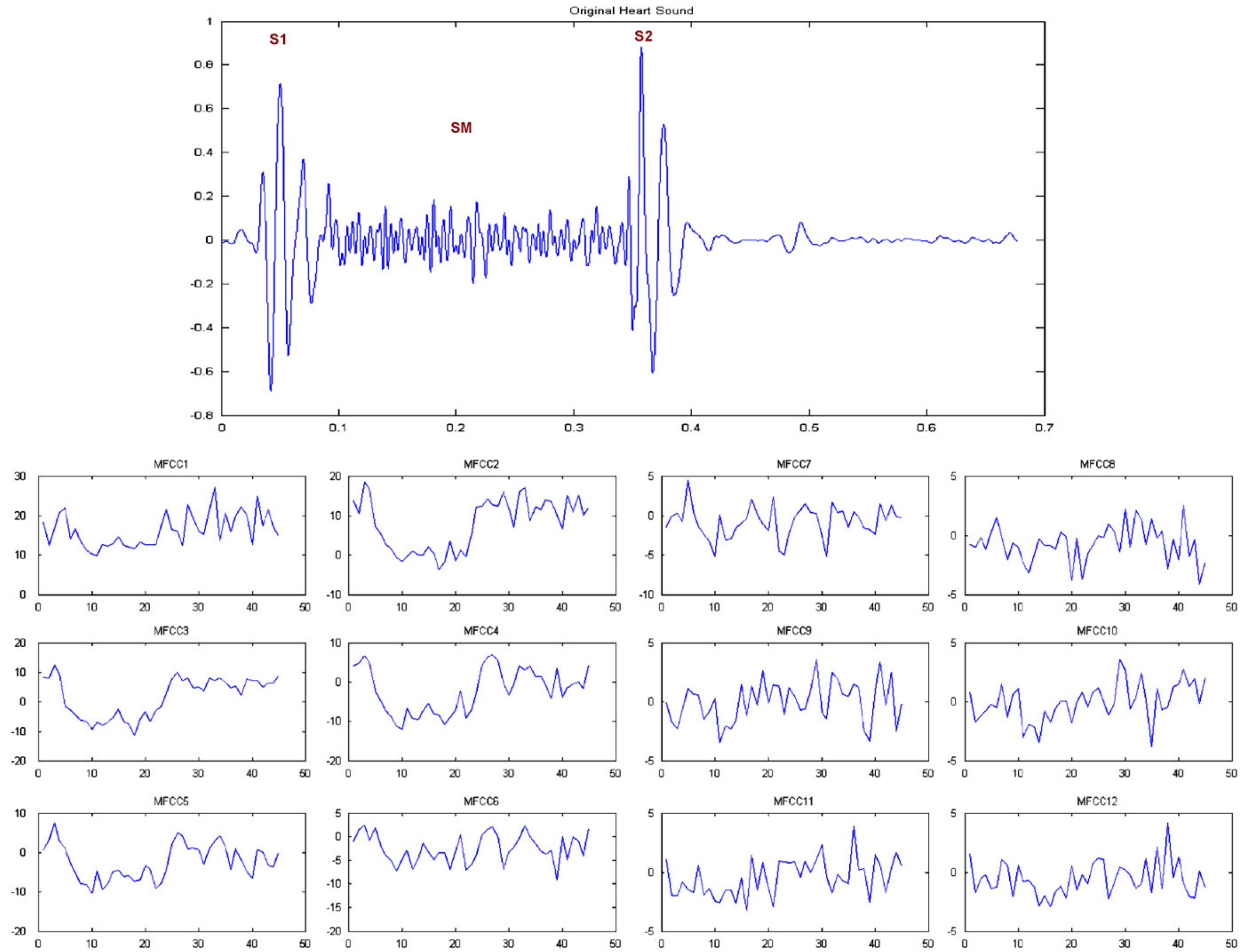


Fig. 7. (a) Heart sounds with systolic murmurs before MFCC and (b) MFCC coefficients.

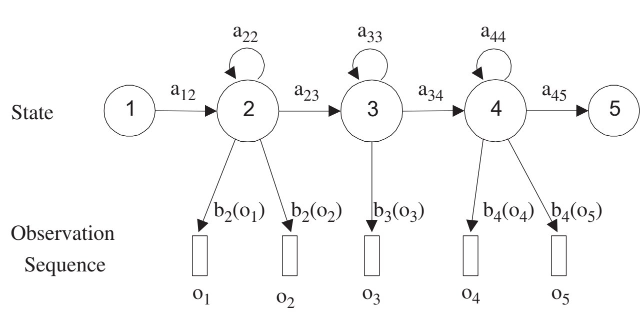


Fig. 8. 5-State hidden Markov model.

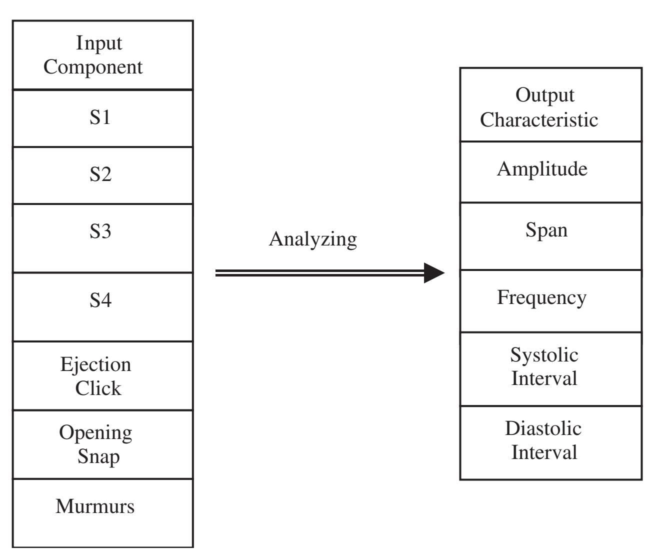


Fig. 9. Heart sounds characterization.

by a normalized Shannon energy as explained in the earlier section. Consequently, the corresponding quantitative parametric characteristics of each component was calculated to assist clinical doctors with objective parameters such as amplitude, frequency, span, intervals, etc.

Fig. 9 illustrates the characterization of heart sounds. Based on classification results, the threshold of Shannon energy of the heart cycle was adjusted until all possible components were identified. Then the characteristics such as amplitude, frequency, span (time–length of the component) and interval (including systolic interval and diastolic interval) were computed.

3. Results and discussions

In order to evaluate the proposed algorithms, a number of experiments were performed using the collected data and public domain data. A comparative experiment was performed to detect the primary heart sound components $S1$ and $S2$ by using four methods: absolute value, squared energy, Shannon entropy and Shannon energy, respectively. For a set of heart sounds which included $S4$ (Fig. 4(a)), four envelopes of the heart signals are shown in Figs. 4(b)–(e). The absolute value (Fig. 4(b)) assumes the same weight for all constituent signals. Their peak values of $S1$, $S2$, $S4$, noise and the corresponding ratios of $S2/S1$, $S4/S1$ and noise$/S1$ of the second cycle are listed in Table 1. The low intensity sounds (such as $S4$) are buried in the squared energy (Fig. 4(c)) by reducing $S4/S1$ from 0.6299 to 0.3968 (Table 1). Although the Shannon entropy (Fig. 4(d)) accentuates the low intensity sounds ($S4/S1$ reaches 0.9950), it also accentuates the effect of low noise (Noise$/S1$ is 0.1142). However, the Shannon energy emphasizes the medium intensity signal so that the three components of $S1$, $S2$ and $S4$ in the signals are easily identified with high values of $S2/S1$ and $S4/S1$, and the low value of Noise$/S1$.

Table 1
Heart sound detection ($S1$, $S2$, $S4$) using comparative analysis of four envelopes

	$S1$	$S2$	$S4$	Noise	$S2/S1$	$S4/S1$	Noise$/S1$
Absolute value	0.9901	0.6175	0.6237	0.0088	0.6237	0.6299	0.0089
Squared energy	0.9803	0.3813	0.3890	0.0041	0.0391	0.3968	0.0077
Shannon entropy	0.3679	0.3679	0.3677	0.0420	1.0000	0.9950	0.1142
Shannon energy	0.3679	0.3679	0.3678	0.0008	1.0000	0.9997	0.0022

The segmentation procedure of heart sound record which included $S4$ is shown in Fig. 5. Shannon energy was calculated and an initial threshold was set to 0.2 times of the maximum to identify $S1$ and $S2$. A limit window with 50 ms was used to scan throughout the whole set of Shannon energy signals as shown in Fig. 5(b). The threshold was increased or reduced, respectively, to reject the extra peaks or to recover the lost peaks. After calculating the threshold, the whole set of normalized average Shannon energy was compared with the threshold to eliminate the effect of noises so as to identify $S1$ and $S2$ (Fig. 5(c)). Based on pause lengths, $S1$ (Fig. 5(d)) and $S2$ (Fig. 5(e)) were identified separately. Thus, raw signals were segmented into cycles (Fig. 5(f)).

A comparative experiment was performed using three methods of feature extraction, viz. the time-domain feature, STFT and MFCC. Twenty sets of testing data were fed into the corresponding system and RCP was used to evaluate the algorithm. Table 2 shows the RCP values obtained. The average RCP of time-domain feature is 93.75%, which indicates that the time-domain feature was not the best representation of the heart sound signals for it contained only time-domain information while information hidden in the frequency content could not be distinguished. The RCP of STFT could reach 97.5%. To extract MFCC features, the heart sound cycle was divided into a number of overlapping frames, each 25 ms long and shifted by 10 ms to avoid suspected data loss. DCT was applied to the logarithm of the filter-bank outputs to calculate the MFCC coefficients. Twelve Mel spectrum coefficients were calculated as shown in Fig. 7. MFCC was obtained by warping the spectrum to the Mel scale which resulted in 100% of the RCP, for it was very similar to perceptual linear predictive analysis of sound and thus optimized the computer-aided auscultation system. The MFCC features of heart cycle with systolic murmurs are shown in Fig. 7. The results showed that the proposed method using MFCC can help in making an effective interpretation.

Four HMM models were trained by using clinical data from 41 samples. Models were trained and compared with different state numbers N and Gaussian mixtures M. Twenty sets of the data were used to test the models. The comparison results are

Table 2
Comparative feature extraction

Number of feature cycles	Type of heart sounds	Correct	Total	Recognize correct percentage (RCP %)	Average RCP (%)
Time domain	CM	20	20	100	93.75
	DM	17	20	85	
	SM	20	20	100	
	Normal	18	20	90	
STFT	CM	19	20	95	97.50
	DM	20	20	100	
	SM	20	20	100	
	Normal	19	20	95	
MFCC	CM	20	20	100	100
	DM	20	20	100	
	SM	20	20	100	
	Normal	20	20	100	

CM, continuous murmurs, DM, diastolic murmurs, SM, systolic murmurs.

Table 3
Recognition results with different states

Models	Cycle number				
	Input test data	Correct	Incorrect	Total	Correct percentage (%)
State 3	CM	16	4	20	80
	DM	19	1	20	95
	Normal	20	0	20	100
	SM	20	0	20	100
State 4	CM	17	3	20	85
	DM	20	0	20	100
	Normal	20	0	20	100
	SM	20	0	20	100
State 5	CM	20	0	20	100
	DM	20	0	20	100
	Normal	20	0	20	100
	SM	20	0	20	100

CM, continuous murmurs, DM, diastolic murmurs, SM, systolic murmurs.

Table 4
Recognition results with different Gaussian mixtures

Models	Cycle number				
	Input test data	Correct	Incorrect	Total	Correct percentage (%)
12	CM	16	4	20	80
	DM	20	0	20	100
	Normal	20	0	20	100
	SM	19	1	20	95
14	CM	18	2	20	90
	DM	19	1	20	95
	Normal	20	0	20	100
	SM	20	0	20	100
16	CM	20	0	20	100
	DM	20	0	20	100
	Normal	20	0	20	100
	SM	20	0	20	100

shown in Tables 3 and 4. Models with 5 states and 16 Gaussian mixtures were found suitable for heart sounds' representation with the maximum possible correctness. Some abnormal heart sound data from public domain [24] were also used so that the proposed algorithms could be tested in a wider domain. In these experiments, 1381 sets of data were fed into the HMM models and 99.21% correct classification was obtained. The testing results are shown in Table 5. The system not only characterizes $S1$ and $S2$ sounds, but abnormal components such as murmurs, $S3$, etc. It also displays the digital characteristics of heart sound components. The amplitude, frequency, span and interval are calculated and displayed to assist physicians for quantitative comparison of relative characteristics.

The main limitation of the present study is the difficulty to obtain in-depth, fast interpretation from the acquired recordings. With the recent rapid development of electronic instruments and signal processing technology, the features and characteristics of heart sounds can be analyzed with appropriate digital signal processing techniques. Only diseases corresponding to known features of heart sounds could be indicated directly. Further studies will be required to investigate the usefulness of the proposed heart sound analysis methods in an environment where multiple sources of loud noise and disturbances can influence the recordings. Secondly, in the segmentation section, the Shannon energy-based rejecting and recovering algorithm achieved good performance in detecting heart sound cycles where $S1$ and $S2$ are the dominating components. However, in the case that murmurs have larger energy in systole and diastole than $S1/S2$, such as some continuous murmurs, the system may detect murmurs as $S1/S2$ by mistake. We are working on the improved algorithms to reduce such fault segmentation.

Table 5
Recognition results of the system

Number of cycles	Correct	Incorrect	Total	Correct percentage (%)
Continuous murmurs	387	4	391	98.98
Diastolic murmurs	144	7	151	95.36
Normal	444	0	444	100
Systolic murmurs	371	0	371	100
Ejection click	5	0	5	100
Opening snap	6	0	6	100
Split $S1$	4	0	4	100
Split $S2$	5	0	5	100
$S3$	9	0	9	100
$S4$	6	0	6	100
Total	1381	11	1393	99.21

Although our computerized system was devised to carry out classification of heart sounds, it can be extended for a knowledge-based classification [25,26] and medical disease database for diagnosis. Nonetheless, there are considerable efforts focused on quantitative and systematic methods for heart sound analyses. This can be due to a statistical requirement of large number for constructing an appropriate general decision tree structure. However, it is understandable that all cardiac diseases may not be interpreted by heart sound analysis alone.

4. Conclusion

This paper presents a computerized system for heart sound signal classification based on a combination of MFCC and HMM, which establishes a preprocessing procedure that includes data acquisition, filtering, normalization and segmentation, followed by the MFCC to efficiently extract the features for the pre-processed heart sound cycles with an automatic procedure that uses HMM. This study advances and investigates various signal processing and classification techniques for an automated cardiac sound classification for helping diagnosis, with the intent of providing a computerized auscultation device that can record and process the sound signals, displays them with visualization playback for classification and characterization. The results show that the system provides an effective aid for parametric and quantitative analyses of auscultatory signals which would help obviating subjective manual diagnosis using traditional stethoscopes. Further studies will be required to investigate the usefulness of the proposed heart sound analysis and methods in an environment where background noises and disturbances can influence the recordings. Moreover, classification algorithms such as neural networks and decision tree could be applied for automated diagnosis.

5. Summary

Auscultation is an important diagnostic means for cardiovascular analysis. This study uses a combination of MFCC and HMM to efficiently extract the features for pre-processed heart sound cycles for the purpose of automatic classification.

In this study, we developed a system for automated interpretation of the PCG signals using pattern recognition techniques. The diagnostic performance of the characterization was demonstrated on the heart sound signals. The task of feature extraction was performed using time features, STFT and MFCC. A comparison of feature extraction was performed and the results showed that the proposed method using MFCC can make an effective interpretation.

Following feature extraction, an automatic classification process was performed by using HMM. Satisfactory classification results were achieved by clinical and downloaded (public domain) data. The RCP varies from 95.36% to 100% with the mean of 99.21%. Furthermore, by the characterization module of this system, various characteristics of constituents of heart sounds were evaluated. Heart sound classification and analysis play a crucial role in the auscultative diagnosis. This system can be used as a clinical analysis tool for doctors to achieve more objective diagnosis and treatment.

Conflict of interest statement

There is no conflict of interest for this publication.

References

[1] A.A. Luisada, From Auscultation to Phonocardiography, Mosby, Saint Louis, 1965.
[2] A. Leatham, Auscultation of the Heart and Phonocardiography, second ed., Churchill Livingstone, London, Edinburgh, 1975.
[3] N.C. Nanda, Doppler Echocardiography, second ed., Lea & Febiger, London, 1993.
[4] T.A.D. Michael, Auscultation of the Heart, McGraw-Hill, New York, 1997.
[5] H. Liang, S. Lukkarinen, I. Hartimo, Heart sound segmentation algorithm based on heart sound envelogram, Comput. Cardiol. 24 (1997) 105–108.
[6] H. Liang, S. Lukkarinen, I. Hartimo, Heart sound segmentation algorithm using wavelet decomposition and reconstruction, EMBS'97, Proceeding of IEEE, Chigaco, USA, 1997, pp. 1630–1633.
[7] B. Erickson, Heart Sounds and Murmurs, A Practical Guide, Mosby-Year Book, Inc., St. Louis, 1997.
[8] P.P. Pinna, E. Piccolo, S. Bartolozzi, V. Fontana, The FFT in the study of the fourth heart sound, IEEE Comput. Cardiol. (1979) 369–372.
[9] A. Djebbari, F.B. Reguig, Short-time Fourier transform analysis of the phonocardiogram signal, The Seventh IEEE International Conference on Electronics, Circuits and Systems, Kaslik, Lebanon, 2000, pp. 844–847.
[10] S.L. Xu, M. Hu, D.C. Zhang, The time–frequency representation of PCG, Industrial technology, Proceedings of the IEEE International Conference, Los Angeles, USA, 1994, pp. 679–681.
[11] L. Khadra, M. Matalgah, B. El-Asir, S. Mawagdeh, The wavelet transform and its application to phonocardiogram signal analysis, Med. Inf. 6 (1991) 271–277.
[12] M. Matalgah, J. Knopp, S. Mawagdeh, Interactive processing method using Gabor wavelet and the wavelet transform for the analysis of phonocardiogram signals, Time–Frequency and Wavelets in Biomedical Signal Processing, IEEE Engineering in Medicine and Biology Society. 1997, pp. 271–304.
[13] H.P. Sava, R. Bedi, J.T.E. Mcdonnell, P.M. Grant, Classification of Carpentimer–Edwards bioprosthesis heart valves using an adaptive single layer perceptron, Proceedings of Engineering in Medicine and Biology Society, Montréal, CA, vol. 1, 1995, pp. 129–130.
[14] J.E. Hebden, J.N. Torry, Neural network and conventional classifiers to distinguish between first and second heart sounds. Artificial Intelligence Methods for Biomedical Data Processing, IEE Colloquium, April, London, vol. 3, 1996, pp. 1–6.
[15] M. Akay, Y.M. Akay, W. Welkowitz, Automated noninvasive detection of coronary artery disease using wavelet-based neural networks, Proceedings of the 16th Annual International Conference of Engineering in Medicine and Biology Society, Baltimore, USA, vol. 1, 1994, pp. A12–A13.
[16] M. Akay, Y.M. Akay, W. Welkowitz, Neural networks for the diagnosis of coronary artery disease, International Joint Conference on Neural Networks, Denver, USA, vol. 2, 1992, pp. 419–424.
[17] P. Wang, C.S. Lim, S. Chauhan, J.Y.A. Foo, V. Anantharaman, Phonocardiographic signal analysis using a modified hidden Markov model, Ann. Biomedical Eng. 35 (3) (2007) 367–374.
[18] S.G. Kahalekar, S.J. Vaidya, A. Shrawagi, PC based phonocardio expert system, Proceedings RC IEEE-EMBS and 14th BMESI, vol. 2, New Delhi, India, 1995, pp. 64–65.
[19] S. Molau, M. Pitz, R. Schluter, H. Ney, Computing Mel-frequency cepstral coefficients on the power spectrum, IEEE International Conference on Acoustics, Speech, and Signal Processing, May 7–11, Utah, USA, 2001, pp. 73–76.
[20] Available from: ⟨http://www.lce.hut.fi/∼vmakine2/mscc.pdf⟩.
[21] L.R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, Proc. IEEE (1989) 257–286.
[22] X. Li, M. Parizeau, R. Plamondon, Training hidden Markov models with multiple observations—a combinatorial method, IEEE Trans. PAMI 22 (2000) 371–377.
[23] S. Young, et al., The HTK Book: HTK Tools and Reference Manuals, Version 2.2, Entropic, 1999.
[24] Available from: ⟨http://www.bcm.tmc.edu/class2000/sims/HeartSounds.html⟩.
[25] S. Pavlopoulos, A. Stasis, E. Loukis, A decision tree-based method for the differential diagnosis of aortic stenosis from mitral regurgitation using heart sounds. BioMedical Engineering OnLine, 3:21. Available from: ⟨http://www.biomedical-engineering-online.com/content/3/1/21⟩, 2004.
[26] T. Leung, P. White, W. Collis, E. Brown, A. Salmon, Classification of heart sounds using time–frequency method and artificial neural networks, in: Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 2, 2000, pp. 988–991.

Sunita Chauhan (PhD in Medical Robotics, Imperial College of Science Technology and Medicine, London, UK) is currently working as an Associate Professor at the Mechatronics and Design Division, School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. Her current research interests include Medical/Surgical Robotics, Computer Assisted and Integrated Surgery, Medical Ultrasound (imaging, therapeutic, surgical), Bio-sensors, Image/sensor processing/interpretation/fusion, etc.

Ping Wang (M.Eng, Nanyang Technological University, Singapore) is currently working as a Research Associate at the BioMedical Engineering Research Centre. Her research interests include pattern recognition, biomedical signal processing and biostatistics.

Chu Sing Lim (PhD, Loughborough University, UK) is currently working as an Associate Professor at the Division of Manufacturing Engineering, School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. He is the director of Biomedical Engineering Research Center. His research interests include Mechanics of Respiratory Management, Biomedical Sensors, Bio-Rapid Prototyping and Tooling.

Venkataraman Anantharaman (MBBS, FRCP (Edin), FAMS, FRCSEd (A&E)) is a Clinical Professor, Chairman of the Division of Ambulatory and Clinical Support Services and Senior Consultant at the Singapore General Hospital, which governs Emergency Cardiac Care, Emergency Trauma Care, Disaster Medicine, Emergency Toxicoloy, Emergency Pre-Hospital Care and Airway Management.