Annals of Biomedical Engineering, Vol. 34, No. 11, November 2006 (© 2006) pp. 1666–1677
DOI: 10.1007/s10439-006-9187-4

Feature Extraction for Systolic Heart Murmur Classification

CHRISTER AHLSTROM,[1,2] PETER HULT,[1,2] PETER RASK,[3] JAN-ERIK KARLSSON,[4] EVA NYLANDER,[5] ULF DAHLSTRÖM,[5] and PER ASK[1,2]

[1]Department of Biomedical Engineering, University Hospital, Linköping University, IMT, SE-581 85, Linköping, Sweden; [2]Biomedical Engineering, Örebro University Hospital, Örebro, Sweden; [3]Department of Clinical Physiology, University Hospital, Örebro, Sweden; [4]Department of Internal Medicine, County Hospital Ryhov, Jönköping, Sweden; and [5]Department of Medicine and Care, Linköping University Hospital, Linköping, Sweden

(*Received 8 March 2006; accepted 22 August 2006; published online: 4 October 2006*)

Abstract—Heart murmurs are often the first signs of pathological changes of the heart valves, and they are usually found during auscultation in the primary health care. Distinguishing a pathological murmur from a physiological murmur is however difficult, why an ''intelligent stethoscope'' with decision support abilities would be of great value. Phonocardiographic signals were acquired from 36 patients with aortic valve stenosis, mitral insufficiency or physiological murmurs, and the data were analyzed with the aim to find a suitable feature subset for automatic classification of heart murmurs. Techniques such as Shannon energy, wavelets, fractal dimensions and recurrence quantification analysis were used to extract 207 features. 157 of these features have not previously been used in heart murmur classification. A multi-domain subset consisting of 14, both old and new, features was derived using Pudil's sequential floating forward selection (SFFS) method. This subset was compared with several single domain feature sets. Using neural network classification, the selected multi-domain subset gave the best results; 86% correct classifications compared to 68% for the first runner-up. In conclusion, the derived feature set was superior to the comparative sets, and seems rather robust to noisy data.

Keywords—Auscultation, Bioacoustics, Feature selection, Heart sounds, Valvular disease.

INTRODUCTION

Cardiac murmurs are often the first sign of pathological changes in the heart valves. Doppler-echocardiography and magnetic resonance imaging are today well established tools in the diagnosis of heart valve disorders, while the classic techniques of auscultation and phonocardiography are playing a diminishing role in modern specialist care. However, in primary or home health care, when deciding who requires special care, auscultation still plays a very important role. For these situations, an ''intelligent stethoscope'' with decision support abilities would be of great value.

Heart murmurs are caused by turbulent blood flow or jet flow impinging on and causing vibration of surrounding tissue. Pathological murmurs are caused by flow through stenosed valves, regurgitant flow through incompetent valves or flow through septal defects. Since diastolic murmurs are mostly pathological, only systolic murmurs are considered in this study.

The research on signal processing of heart sound (HS) recordings has been extensive.[4] Several authors have investigated the possibility to automatically classify cardiac murmurs.[2,3,8,9,14,15,21,26,31–33,38,41]. Common for all classification tasks is the importance of appropriate data representations (features). These features should retain similarities within classes while revealing differences between classes. HS are characterized by their timing, morphology and frequency.[37] Suitable features for classification of systolic murmurs should hence be able to describe information in these domains. The feature sets used in previous works often assumes linearity and ranges from time domain characteristics[2,26,32] via spectral characteristics[3,33,41] and frequency representations with time resolution[8,14,15,24,38,41] to parametric modeling.[22,35]

The assumption of linearity basically requires all significant information to be contained in the frequency spectrum. From a stochastic process perspective, power spectral information is described by first and second order statistics. However, heart sounds contain non-linear and non-Gaussian information that is not revealed in the frequency spectra.[5,9,21] Higher order statistics (HOS) is thus motivated. Taking a deterministic viewpoint, dynamical systems theory can be used to describe nonlinear behaviour. A topological equivalent to a systems true state space can be reconstructed with the method of delays.[11] Features derived

Address correspondence to Christer Ahlstrom, Department of Biomedical Engineering, University Hospital, Linköping University, IMT, SE-581 85, Linköping, Sweden. Electronic mail: christer@imt.liu.se

from HOS and from the reconstructed state space will here be investigated for heart murmur classification.

The aims of this study are to:

(a) Develop new features, mostly inspired by research in speech processing and in dynamical systems and chaos theory.
(b) Present a set of features, based on a combination of old and new features, suitable for classification of systolic heart murmurs.
(c) Compare the classification performance of different feature sets.

METHODOLOGY

This section provides a short description of the data, the acquisition method and the patients. This is followed by a survey of different feature extraction methods, reviewing previously used features and introducing new ones.

Data Acquisition and Patients

HS data were recorded at the Dept. of Internal Medicine at Ryhov County Hospital, Jönköping, Sweden and at the Department of Clinical Physiology, University Hospital, Örebro, Sweden. This study was approved by the ethical committee at Linköping University Hospital and all patients enrolled gave their informed consent.

Patients with probable valvular heart disease (as detected with auscultation) were asked to participate in the study. The patients underwent an echocardiographic examination, where diagnosis and severity of valve lesions were determined by experienced echocardiographers according to clinical routine and recommended standards.[30] HS were acquired in association with this examination. In total, 36 patients (19 male, 17 female, ages 69 ± 14 years, all with native heart valves) were enrolled in the study. Based on the results from the clinical echocardiographic examination, an independent physician re-evaluated the echocardiographic reports. According to the two physicians' evaluations, who agreed for all diagnoses, the patients were divided into three groups (6 patients with moderate to severe mitral insufficiency (MI), 23 patients with mild to severe aortic stenosis (AS) and 7 patients with physiological murmurs (PM).

An electronic stethoscope (theStethoscope, Meditron AS, Oslo, Norway) was used to acquire the HS and a standard 3-lead ECG (Analyzer ECG, Meditron AS, Oslo, Norway) was recorded in parallel as a time reference. Both signals were digitized at 44.1 kHz with 16-bits per sample using a sound card (Analyzer, Meditron AS). HS data were recorded successively for 15 s from the four traditional areas of auscultation.[37] Based on signal quality, one of the four signals was selected after visual and auditive inspection. The diagnosis of the patient was not known during the selection process. All processing of the signals was performed in MATLAB (The MathWorks, Inc., Natick, MA, USA).

Features

Automatic extraction of features depends on accurate knowledge about the timing of the heart cycles. Segmentation into the first heart sound (S1), systole, the second heart sound (S2) and diastole is thus needed. A reliable way to do this is by ECG gating, i.e., to look for S1 in a certain time window after the R-peak and for S2 in a time window after the T-wave of the ECG. The local maximum of the HS signal's envelope (calculated by Shannon energy[16]) within each time window was determined as S1 and S2, respectively. The first local minima before and after S1 and S2 was used to determine the boundaries of the heart sounds. The region of interest in this study, focusing on the systolic period, was defined as the start of S1 to the end of S2. The Pan-Tompkins algorithm was employed to find the R-peaks of the ECG, and a simple threshold was used to find the T-wave in each heart cycle.[25] All time instances were checked and corrected manually to avoid timing errors at this stage.

The HS signal will be denoted $s(n)$, where $n = 1, 2,..., N$ and N is the number of samples in the time period from the start of S1 to the end of S2 (with one exception in the calculation of the Gaussian mixture model features, where N is the number of samples in one patient). An example HS signal from a patient with AS can be found in Fig. 1a.

The feature extraction process extracts 207 scalar values per heart cycle (again with one exception in the calculation of the Gaussian mixture model features). Each of these was averaged over all available heart cycles, resulting in 207 features per patient. All features were also normalized to zero mean and unit standard deviation. The calculations behind each feature are explained in detail below and a summary of the features is given in Table 1.

Time Domain Features

The envelope of $s(n)$ was extracted with the normalized average Shannon energy,[16] see Eq. (1). The HS signal was divided into short overlapping segments of 40 ms duration (20 ms overlap), and the Shannon energy was calculated in each segment to obtain time resolution.

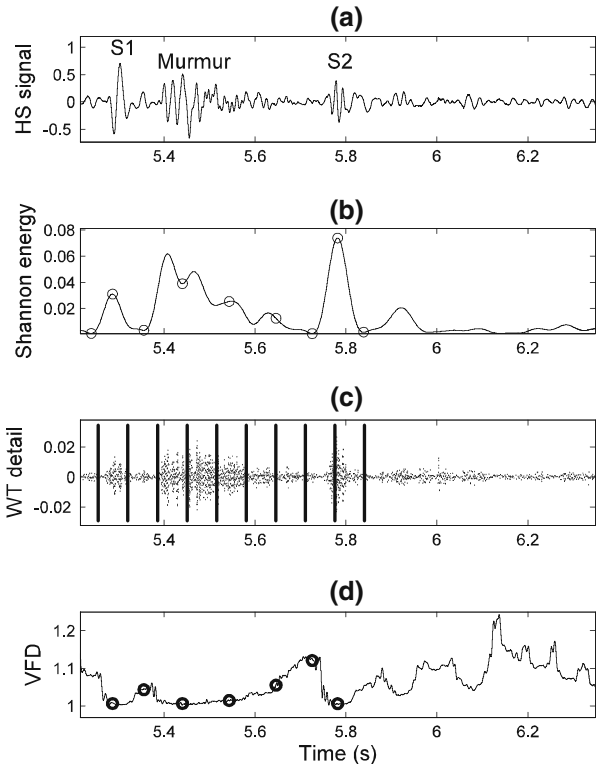

FIGURE 1. An example of a HS signal from a patient with aortic stenosis is shown in (a). In (b) the signal's envelope has been extracted (Shannon energy), and 9 envelope values are chosen as features, *Shannon energy 1–9*. The 6th wavelet detail is illustrated in (c), where the vertical lines are time markers equidistantly distributed over the region of interest. The absolute sum between each marker constituted the feature values, *WT detail 1–9*. In (d) the variance fractal dimension trajectory is plotted together with seven of the fractal dimension features, *VFD 1–7*. The material explaining (c, d) will be covered in following sections. All units are arbitrary.

$$E_{\text{Shannon}} = -\frac{1}{N_{\text{seg}}}\sum_{n=1}^{N_{\text{seg}}} s^2(n)\cdot \log s^2(n) \qquad (1)$$

where N_{seg} is the number of samples in the 40 ms segment. Nine envelope values (denoted *Shannon energy 1–9*) were selected as features; at times before S1, peak S1, after S1, ¼ into systole, ½ into systole, ¾ into systole, before S2, peak S2 and finally after S2. The systolic features were chosen equidistantly between the end of S1 and the start of S2, and the features related to S1 and S2 were selected as local minima and maxima, see Fig. 1b. Similar time domain features have previously been used in other studies.[17,26,32]

Time Frequency Representation (TFR) Based Features

The wavelet transform (WT) and the S-transform[34] (ST) were used to extract features describing how the frequency content of the signal varied over time. WT and ST are defined as:

$$\text{WT}(m,k) = \frac{1}{\sqrt{|k|}}\sum_{n=1}^{N} s(n)w\left(\frac{n-m}{k}\right) \qquad (2)$$

$$\text{ST}(m,k) = \frac{|k|}{\sqrt{2\pi}}\sum_{n=1}^{N} s(n)e^{-\frac{(n-m)^2k^2}{2}}e^{-2\pi i\frac{kn}{N}} \qquad (3)$$

where m is the translation parameter, k is the scale (WT) or frequency (ST) parameter and w is the mother wavelet. The resulting TFR matrices are of considerable size, so data reduction is required to obtain manageable feature sets. Reduction into 16 features was achieved by limiting the frequency content of the S-transform to 150 Hz and down-sampling the result into a 4×4 matrix,[14] see Fig. 2. About 150 Hz was chosen as a compromise between the gains in using higher frequencies versus keeping the number of features low. These features were denoted *ST map 1–16*. Furthermore, one of the wavelet details was down-sampled into a number of features.[8,24] Here the 6th wavelet detail of a level 10 Daubechies 2 wavelet decomposition was discretized into nine features (*WT detail 1–9*) by taking the absolute sum over equidistantly spaced intervals spanning the systolic period, see Fig. 1c (the 6th detail roughly corresponds to the pseudo-frequency 1 kHz). The entropy of the wavelet approximation and of each wavelet detail, still from the level 10 Daubechies 2 wavelet, were also used as features, *WT entropy 1–11*.

A perhaps more refined method for data reduction of TFR matrices use singular value decomposition, $\text{TFR} = U\Sigma V^T$. U and V are generally called left and right eigenvectors, or in this particular case eigentime and eigenfrequency, respectively. To create a compact representation, a distribution function was extracted for eigentimes and eigenfrequencies corresponding to the two largest eigenvalues (since eigenvectors are orthonormal, their squared elements can be considered as density functions[10]). A histogram (10 bins) was computed for each distribution function leading to 40 new features (*Eigenfrequency 1 1–10, Eigenfrequency 2 1–10, Eigentime 1 1–10 and Eigentime 2 1–10*). The first eight eigenvalues were also used as features (*Eigenvalue 1–8*). In the singular value decomposition calculations, the TFR was derived using the S-transform. The TFR, the eigenvalues, the eigenvectors, the probability distribution functions and the histograms are illustrated in Fig. 3. These features can be interpreted as the main components of the TFR matrix. For example, the minima of the first eigentime in Fig. 3i correspond to S1, S2 and the murmur.

TABLE 1. Summary of all features.

Time domain features	*Shannon energy 1–9*	9	Envelope values derived from the normalized Shannon energy
Time frequency relation based features	*WT entropy 1–11*	11*	The Shannon entropy of each wavelet detail and the wavelet approximation using a level 10 decomposition with the Daubechies 2 wavelet
	WT detail 1–9	9	The 6^{th} wavelet detail of a level 10 Daubechies 2 wavelet discretized into 9 bins
	ST map 1–16	16	TFR (calculated with the S-transform) in the frequency range 0–150 Hz discretized into a 4×4-matrix
	Eigenvalue 1–8	8*	The eight first eigenvalues from a singular value decomposition of an S-transform TFR
	Eigenfrequency 1 1–10, Eigenfrequency 2 1–10, Eigentime 1 1–10, igentime 2 1–10	40*	Two left eigenvectors and two right eigenvectors transformed into distribution functions whose histograms (10 bins) are used as features
Nonlinear and chaos based features	*HOS 1–16*	16*	First non-redundant region of the bispectrum (frequency range 0–300 Hz) discretized into 16 equally sized triangles
	GMMx cycle 1–8	40*	Gaussian mixture Model (GMM) of the reconstructed state space of the systolic period (including HS). *x = 1, 2,..., 5* represents the mixtures and *1–8* represents four coordinates and four eigenvalues
	GMMx murmur 1–8	40*	GMM of the reconstructed state space of the systolic period (excluding HS). Notation as above
	VFD 1–8	8*	Variance fractal dimension values
	RQA 1–10	10*	Recurrence Quantification Analysis

Column one through four represents the category of the extraction method, the feature names, the number of features and a short description of the extraction technique, respectively.
*Features not previously used for heart murmur classification.

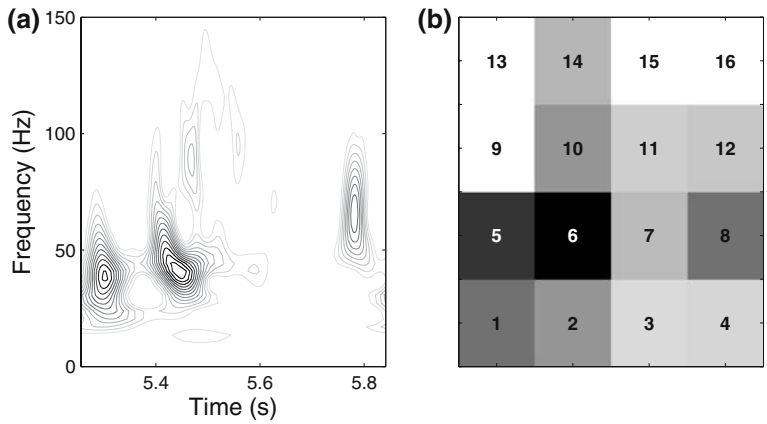

FIGURE 2. Time frequency representation (calculated with the S-transform) of one heart cycle from a patient with aortic stenosis (a), S1 can be seen at 5.3 s and S2 at 5.8 s. The greyscale represent intensity where darker shades represent higher intensities. In (b) the same data has been discretized into a 4×4 map of features, where the numbers represent the features denoted ST map 1–16.

Nonlinear and Chaos Based Features

A statistical approach commonly used to analyze nonlinear signals is HOS, while a deterministic viewpoint often leads to the method of delays (a tool to reconstruct the state-space geometry of a dynamic system using only the observations of a single component). The method of delays provides a foundation for many analysis methods, ranging from graphical representations to the calculation of numerical characteristics.

Higher Order Statistics

Higher order statistics preserve the phase character of signals, and can be used to describe nonlinear or non-Gaussian processes.[9] For truly Gaussian processes, all cumulant spectra of order greater than two

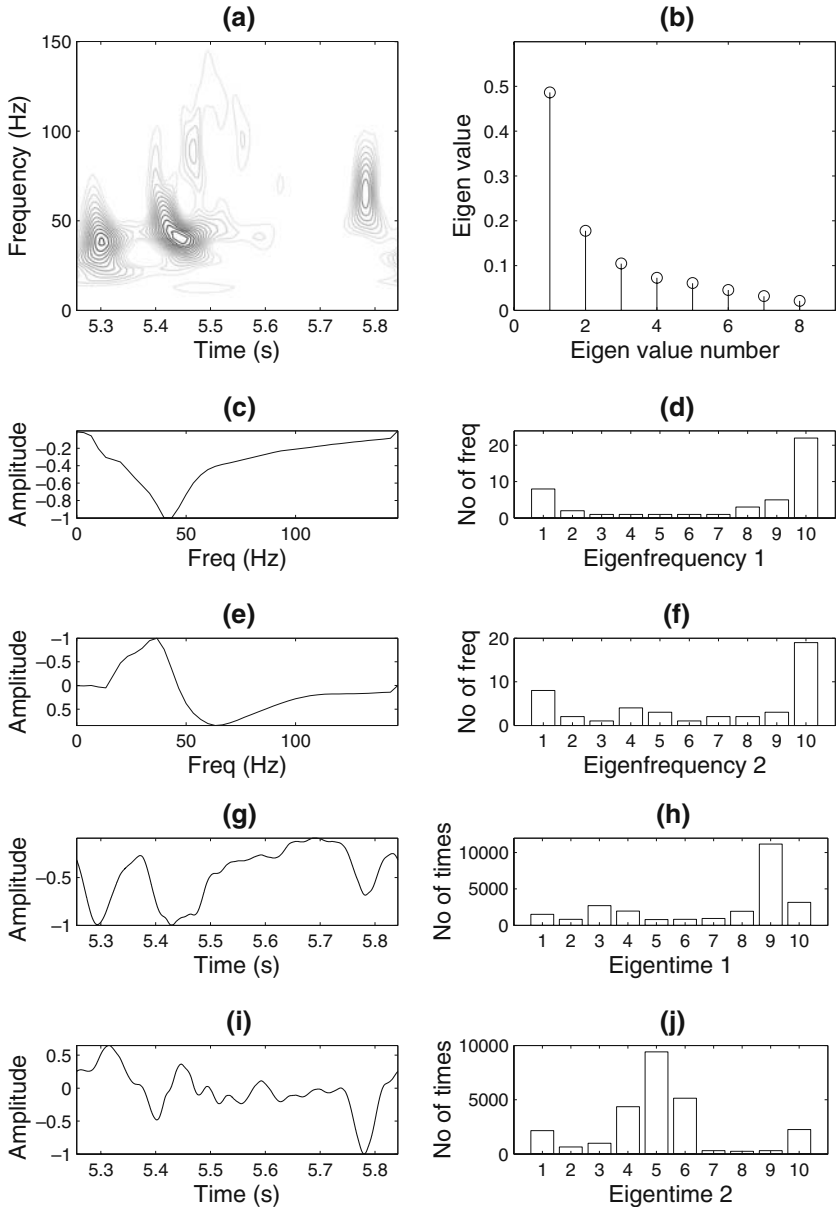

FIGURE 3. The TFR from Fig. 2 is shown in (a) and *Eigenvalue 1–8* in (b). Part (c) and (e) of the figure illustrates the 1st and 2nd eigenfrequencies. To the right of the respective figures are the histograms of the probability distributions belonging to the eigenfrequencies, *Eigenfrequency 1 1–10* and *Eigenfrequency 2 1–10* (d, f). ''No of frequencies'' is a measure of how significant various parts of the distribution functions are. Parts (g-j) of the figure show corresponding plots for the 1st and 2nd eigentimes (with the features *Eigentime 1 1–10* and *Eigentime 2 1–10*). All units are arbitrary unless stated otherwise.

are zero. HS data have a non-zero bispectrum (the Fourier transform of the third order cumulant), and this information may be used for classification. The bispectrum is defined as:

$$C(k_1,k_2)=\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}E\{s(n)s(n+n_1)s(n+n_2)\}\cdot e^{-2\pi i\frac{k_1n_1+k_2n_2}{N}} \quad (4)$$

where n_1 and n_2 are two lag variables, k_1 and k_2 are the frequency variables and E is the expectation operator. For computational reasons, each heart cycle was downsampled with a variable factor of about twelve so the number of samples ended up below 2000 points (this corresponds to a sampling frequency of about 3500 Hz, thus leaving most of the information in the phonocardiography data intact). The downsampled signal was zero-padded to 2048 samples, partly for

computational efficiency in the FFT calculations, but also for simplicity since all matrices ended up with the same size (2048×2048 values). The bispectral estimates were averaged across records, giving one output bispectrum per patient. Due to symmetry, see Fig. 4, only the first non-redundant region was used for feature extraction (*HOS 1–16*). Bispectra have previously been used to visualize and analyze HS data[5,9] and the feature extraction approach was inspired by Xiang and Tso,[42] who used it to classify flaws in concrete structures. The HOSA Toolbox was used for the calculation of bispectra.

State Space

An observed signal is only a projection of a systems multivariate state space onto a one-dimensional time-series. However, the state space can be reconstructed, at least to a topological equivalent, with Takens' delay embedding theorem[11]:

$$a(n) = \{s(n), s(n+\tau), s(n+2\tau), ..., s(n+(d-1)\tau)\} \tag{5}$$

The embedding parameters, d and τ, was estimated using Cao's[1] method and the average mutual information technique[11], respectively. The embedding dimension was determined as $d = 4$, see the clearly defined knee in Fig. 5b. The time delay was set to 150 since roughly half of the patients had a minimum in the vicinity of $\tau = 150$. The other half lacked minima in the range $\tau = \{1, 2, ..., 400\}$ samples.

In the reconstructed state space, a vector connects the states, thus creating a trajectory describing how the system evolves over time. In this study an estimate of the distribution of the trajectory was used to describe its behaviour. A Gaussian mixture model (GMM) with five mixtures, see Fig. 6, was fitted to the reconstructed 4D state space using the Expectation-Maximization (EM) algorithm. The centres of the mixtures and the eigenvalues of the covariance matrices constituted 40 new features (4 coordinates and 4 eigenvalues for each of the five mixtures). Data from all heart cycles for each patient were used to reconstruct the state space, and two sets of features were calculated based on data ranging from either the beginning of S1 to the end of S2 (*GMMx cycle 1–8*) or from the end of S1 to the beginning of S2 (*GMMx murmur 1–8*). *x = 1, 2,..., 5* represents the five mixtures and *1–8* represents the four coordinates (1–4) and the four eigenvalues (5–8). This approach is based on a speech classification algorithm.[28] The nonlinear time series analysis was performed with TSTool[20] and the GMM of the embedded state space was calculated with the TKDE toolbox.[28]

Fractal Dimension

A signal's complexity, in terms of morphology, entropy, spectrum and variance, can be described by its fractal dimension.[12] Comparing HS, murmurs and background noise, HS have a certain structure while murmurs are more complex and noise has no structure at all.[23] This is all reflected in the fractal dimension. The variance fractal dimension (VFD) was used to estimate the fractal dimension. The VFD for a 1D time series is calculated via the Hurst exponent as $VFD = 2-H$, where H is defined as:

$$H = \lim_{(n_2-n_1)\to 0} \left(\frac{\log[\mathrm{Var}(s(n_2 - n_1))]}{2 \cdot \log[n_2 - n_1]} \right) \tag{6}$$

In practise, a log-log plot is created for a series of dyadic time increments, n_2-n_1, and the slope of the regression line determines H. The HS signal was

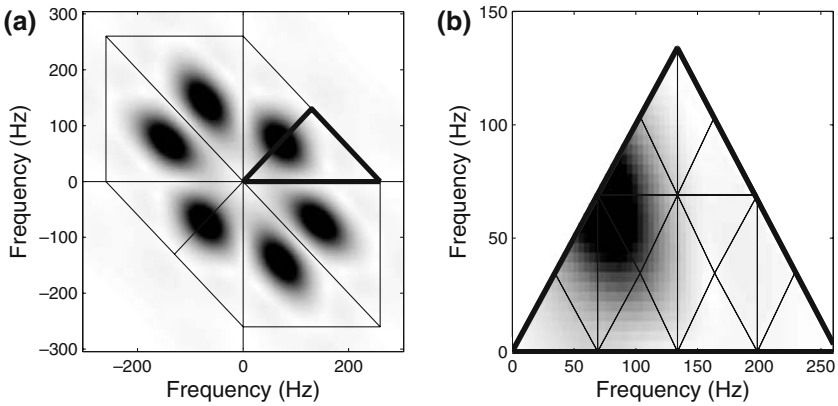

FIGURE 4. Example of bispectrum from a patient with aortic stenosis. The different regions of the bispectrum are plotted in (a) where the bold triangle shows the non-redundant information. In (b) the region of interest is highlighted. The smaller triangles indicate the features *HOS 1–16*.

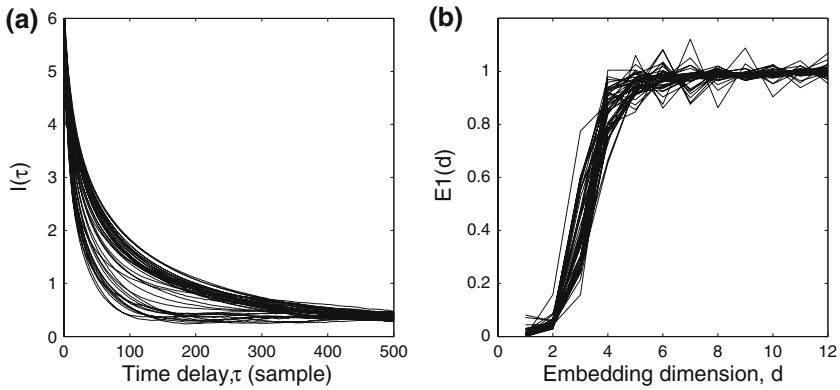

FIGURE 5. The average mutual information, *I*(τ), calculated for each patient is shown in (a), revealing a delay parameter of about τ = 150. Cao's method was used to calculate the E1 measure for each patient in order to determine the embedding dimension to *d = 4* (b). *I* is measured in bits and *E1* is measured in arbitrary units.

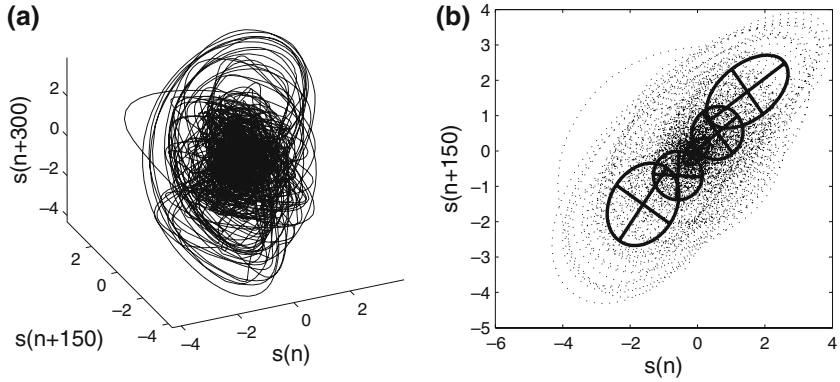

FIGURE 6. The reconstructed state space (*d = 3*, τ = 150) of the systolic period from a patient with aortic stenosis (a). In (b) the same sound, reconstructed with *d = 2*, is shown together with the derived Gaussian mixture model (*GMMx murmur 1–8*). The principle axes of the ellipses indicate one standard deviation of each mixture in the model. Note that *d = 4* is required to unfold the trajectory.

divided into short overlapping segments of 40 ms duration (20 ms overlap), and the VFD was calculated in each segment to obtain time resolution, thus creating a VFD trajectory, see Fig. 1d. More details about VFD trajectories can be found in Kinsner[12,13]. Seven VFD values along the trajectory were selected as features, *VFD 1–7*. These were chosen at S1, after S1, ¼ into systole, ½ into systole, $\frac{3}{4}$ into systole, before S2 and at S2. The systolic features were chosen equidistantly between the end of S1 and the start of S2, and the features related to S1 and S2 were selected as local minima and maxima, see Fig. 1d. The quotient between the minima of S1 and S2 and the minima of the five systolic VFD values was also used as a feature, *VFD 8*. Fractal dimension trajectories have previously been used to locate S1 and S2,[7,23] but to our knowledge not to classify heart murmurs.

Recurrence Quantification Analysis

A recurrence plot (RP) is a binary $N \times N$ matrix representing the recurrence of states of a system (i.e., how often a small region in state space is visited). If two states on the trajectory, *a(i)* and *a(j)*, are close to each other, a black dot is placed at position *(i,j)* in the RP matrix according to:

$$\mathrm{RP}(i,j) = \Theta(\varepsilon - \|a(i) - a(j)\|) \tag{7}$$

where $i,j = 1,\ldots,N$, ϵ is a cutoff distance, $\|\cdot\|$ is the Euclidian norm and $\Theta(\cdot)$ is the Heaviside function. In this study, ϵ was set to 0.1. An example RP is shown in Fig. 7. RPs can be quantified with recurrence quantification analysis (RQA), and the results from such an analysis were used as features. Recurrence plots and RQA have, to our knowledge, not previously been

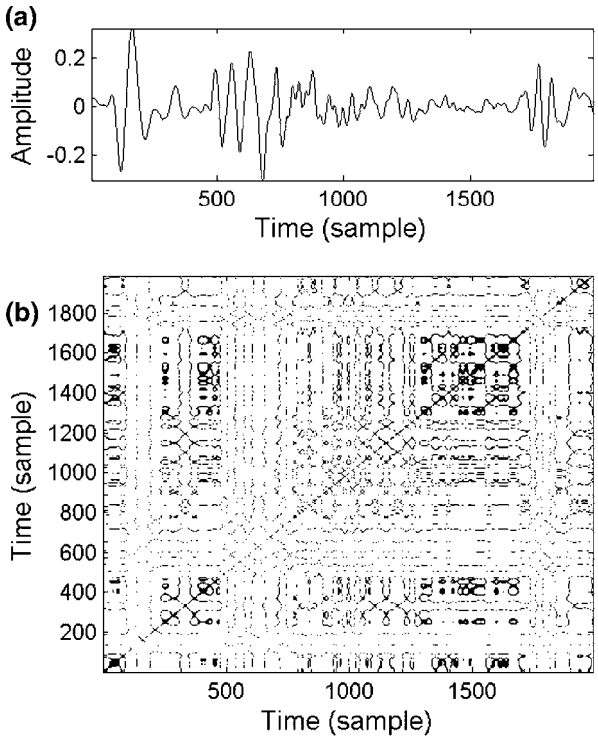

FIGURE 7. An example showing the systolic period in a patient with aortic stenosis (a). The recurrence plot, from where the 10 recurrence statistics are calculated, is shown in (b).

used to classify heart murmurs. The RQA analysis was performed with the CRP toolbox.[19] In this study, the following ten measures were used as features:[6,19,40]

RQA1 Recurrence rate, the percentage of the RP that is filled with recurrence points. This measure corresponds to the correlation sum.

RQA2 Determinism, the percentage of the recurrence points forming diagonal lines (diagonal lines are associated with deterministic patterns in the dynamics).

RQA3 The average length of the diagonal lines, which is related to the predictability of the dynamical system.

RQA4 The length of the longest diagonal line (which is inversely proportional to the largest Lyapunov exponent which describes how fast trajectories diverge in the reconstructed state space).

RQA5 Entropy, the Shannon entropy of the distribution of the diagonal line lengths (measures the complexity of the deterministic structure in the system).

RQA6 Laminarity, the percentage of recurrence points that forms vertical lines. Laminarity relates to the amount of laminar states in the system.

RQA7 Trapping time, the average length of the vertical lines. The trapping time contains information about the frequency of the laminar states and their lengths.

RQA8 The length of the longest vertical line (reveals information about the time duration of the laminar states).

RQA9 Recurrence time of the first kind, relates to the information dimension.

RQA10 Recurrence time of the second kind, relates to the information dimension.

Feature Selection

A large number of features have been described in previous sections. However, too many features often result in higher computational complexity, mutually correlated features and classifiers with low generality.[36] In our study, Pudil's sequential floating forward selection (SFFS) method was used to reduce the number of features.[29] Inclusion or rejection of features was based on the error estimate of a 1-nearest neighbour leave-one-out classifier where the estimation error was used as performance criterion. The resulting subset of features was denoted the SFFS subset.

Classification Performance

A neural network was adopted to measure the performance of the SFFS subset. For comparison, each methodology was also tested separately, i.e., the eleven feature subsets constituted by *Shannon Energy 1–9, WT entropy 1–11, WT detail 1–9, ST map 1–16, Eigenvalue 1–8, Eigentime + Eigenfrequency, HOS 1–16, GMMx cycle 1–8, GMMx murmur 1–8, VFD 1–8 and RQA 1–10*. Each neural network was a fully

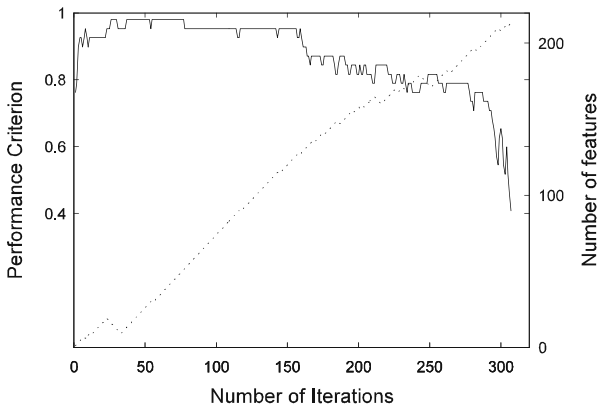

FIGURE 8. The evolution of Pudil's sequential floating forward selection algorithm. The solid line indicates classification performance while the dotted line indicates the number of features in the present feature subset. The feature set with as few features as possible is chosen under the condition that the performance criterion is maximized.

connected feed-forward network, with logarithmic sigmoid transfer functions and biased values throughout. The number of input units was set to the nearest larger integer of the square root of the number of features in the set, the number of units in the hidden layer was set to three and the number of output units was set to two. The target values were 00 (MI), 01 (AS) or 10 (PM). Each output from the network was thresholded at 0.5 and compared to the results from the clinical echocardiography investigation. A leave-one-out approach was used for training and testing due to the limited amount of patients.

RESULTS

The studied material consisted of 445 heart cycles from 36 patients, or 12.4 ± 4.0 (SD) heart cycles per patient.

The number of features in the subset, when maximizing the performance criterion while keeping the number of features low, was 14, see Fig. 8. Nine of these 14 features have not previously been used to classify HS. Table 2 presents *the SSFS subset*.

TABLE 2. The 14 features of the SFFS subset selected with Pudil's sequential floating forward selection method.

1.WT detail 7	2. VFD 8*
3. Shannon energy 5	4. Shannon energy 6
5. GMM1 cycle 5*	6. Shannon energy 4
7. GMM1 murmur 5*	8. Eigenfrequency-1 2*
9. WT entropy 10*	10. GMM4 cycle 6*
11. Eigenfrequency-1 1*	12. Shannon energy 8
13. VFD 2*	14. HOS 1*

The features are ordered in correspondence to classification performance. Features previously not used for murmur classification are denoted with stars.

Confusion matrices showing classification results for the twelve tested subsets are presented in Table 3. The percentage of correct classifications is summarized in Fig. 9 together with the number of patients with valve pathology that were erroneously classified as physiological. The SFFS subset gives the best classification results while the VFD technique provides the best single-domain subset.

DISCUSSION

The main tasks for the intelligent stethoscope are to detect abnormal events (such as the third heart sound or the reverse splitting of S2) and, regarding decision support, to classify different heart valve diseases and distinguish between pathological and physiological murmurs. A feature set making use of HS characteristics from several signal domains has been derived. The presented feature set is able to differentiate between MI, AS and PM with 86% correct classifications. The derived feature set aims at facilitating the classification step by providing relevant information to the classifier. Since diastolic murmurs are mostly pathological,[35] this study was limited to systolic murmurs.

The non-linear features used in this study are not easy to interpret. When leaving the well-known concepts of time and frequency, the obtained features become hard to explain in terms of physiological events. Another complicating issue is that the reconstructed state space is four-dimensional, making it impossible to visualize. A most welcome exception is the fractal dimension measurements since HS, murmurs and background noise can be described by various degrees of complexity (HS have a certain structure while murmurs are more complex and noise has no structure at all). Both RQA

TABLE 3. Confusion matrices showing the classification results from the different feature subsets.

	Shannon Energy			WT entropy			WT detail			ST map		
	AS	MI	PM	AS	MI	PM	AS	MI	PM	AS	MI	PM
AS	17	3	3	14	7	2	15	6	2	14	8	1
MI	4	2	0	3	2	1	5	0	1	4	1	1
PM	4	1	2	5	1	1	2	4	1	3	0	4
	Eigenvalue			Eigentime & freq			HOS			GMMx cycle		
	AS	MI	PM	AS	MI	PM	AS	MI	PM	AS	MI	PM
AS	15	2	6	18	4	1	14	8	1	13	3	7
MI	4	0	2	6	0	0	4	1	1	5	1	0
PM	3	1	3	3	2	2	4	1	2	4	1	2
	GMMx murmur			VFD			RQA			SFFS		
	AS	MI	PM	AS	MI	PM	AS	MI	PM	AS	MI	PM
AS	15	7	1	20	2	1	8	8	7	19	2	2
MI	3	2	1	4	2	0	3	1	2	1	5	0
PM	6	0	1	5	0	2	0	3	4	0	0	7

Target groups are presented horizontally while the predicted groups are presented vertically.

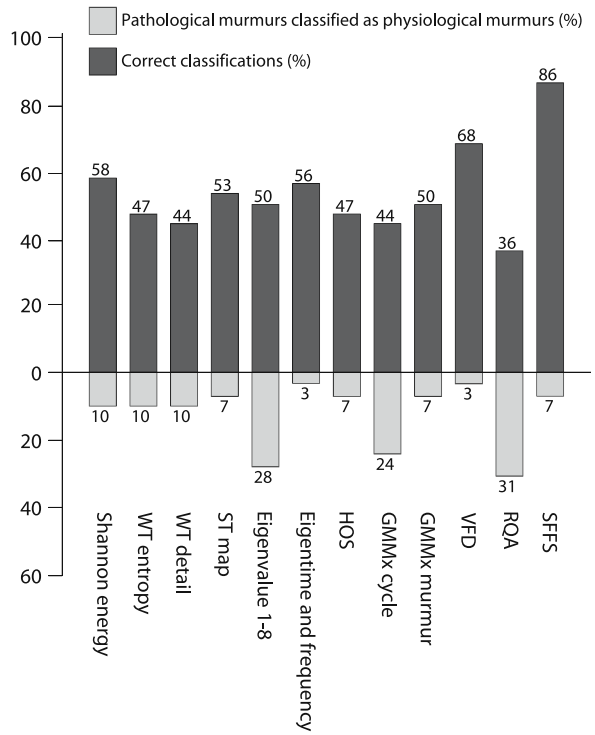

FIGURE 9. Bar graph showing the number of correct classifications for each feature subset when used as input variables to the neural network. Also presented is the number of cases where pathological murmurs are erroneously classified as physiological. The number attached to each bar represent the exact height of the bar in percent.

and the GMM of the reconstructed state space aims at quantifying the amount of structure in the signal. The techniques are potentially useful for HS studies but, because of the pure data-analysis character of the methods, they do not allow for any physiological speculation as such. The selected SFFS subset does however include very reasonable features. *WT detail 7* is located at the end of systole and is thus able to separate the holosystolic MI from diamond-shaped AS and PM. *Shannon energy 4–6* are located in mid systole and describes the shape of the murmur in the time domain. *VFD 8* describes the complexity of the murmur in relation to the heart sounds and could partly be explained as the intensity of the murmur. *VFD 2* is located at S1, which is known to decrease in intensity during MI. *GMM1 cycle 5* and *GMM murmur 5* belong to the largest mixture model, probably located in their respective state spaces where the murmur lives. The selected features of the GMM (the first eigenvalue), describes the size or width of the Gaussian distribution, and the two features should hence relate the size of the murmur to the size of the heart sounds. *Eigenfrequency-1 1&2* contains the low frequency content of the main frequency component, which gives good coverage of the low frequency content of the signal.

Some of the subsets chosen for comparison are very similar to feature sets used in other studies. *Wavelet detail 1–9* was adopted from Gupta *et al.*[8] and Olmez *et al.*[24] and *ST map 1–16* was adopted from Leung *et al.*[14]. *Shannon energy 1–9* was included since these features give a very appealing time domain interpretation. When including features already known from the literature, a nearly full coverage was aimed at. Adaptations of the previously used methods were however necessary. In Gupta *et al.*[8] and Olmez *et al.*[24], a level 6 Daubechies wavelet was used, but to account for the higher sampling frequency in this study, a level 10 wavelet decomposition was used instead. We also used 9 features spread evenly over systole instead of the 32 features spread over the whole heart cycle. A common remark when using WT is that the link to local frequency is lost (why the term scale is preferred instead of frequency). The ST is a similar but phase corrected transform, able to maintain the notion of frequency. In the calculation of the 4×4 map of the TFR, the ST was used instead of WT as in Leung *et al.*[14] and the frequency content was limited to 150 Hz compared to 62 Hz. The ST has previously been used for visual analysis of HS signals.[18,39]

There is a tendency in several methods to classify MI and PM as AS. The total number of MI + PM patients is 13, and out of these patients, 9 are classified incorrectly as AS using VFD, GMMx cycle, GMMx murmur and Eigentime/freq, 8 are classified incorrectly as AS using Shannon energy, WT entropy and HOS while 7 are classified incorrectly as AS using ST map, WT detail, and Eigenvalue. Many of the features within each feature set are similar despite being derived from different diseases, while only a few of the features within the feature set contain the information needed to distinguish the different diseases. It is thus easy to confuse the classifier with the obscure content of non-significant features. In the SFFS subset, only the most descriptive features are used, and it is not surprising that the error rates decreases.

All time domain features used in the literature could not be used. For example, the splitting parameter of the second HS used in Liang *et al.*[15] could not be automatically extracted using the described wavelet method. A considerably more complex transient chirp model has been used to perform the same task[27] (extracting the aortic and pulmonary components of S2), but despite the more advanced methodology, the authors needed multiple sensors to obtain robust results. Since only one sensor was used in our study, this approach was not considered. Using multiple sensors to record simultaneous HS at several locations is however an interesting idea. The magnitude of different components in the phonocardiographic signal varies with the measurement location. For instance,

listening over the apex, S1 is louder than S2. Also, the location where a heart murmur is best heard often indicates its origin. By using multiple sensors in parallel, this difference in intensity could be used as a parameter in a classification system. Incorporating such information in the classification system is left for future studies.

If too many features are used, the performance of the classifier will decrease with respect to execution time (due to the measurement cost) as well as recognition rate (due to overfitting). The reason lies in the existence of many different solutions that are consistent with the training examples, but disagree on unseen (test) examples. This is probably what happens in Fig. 8, where the performance starts going down when more than about 150 features are added.

Previous studies present excellent classification results well above 95 % when classifying a number of different heart abnormalities.[2,3,8,14,24,26,32,33,38,41] In this study the number of successful classifications was at best 86% when classifying MI, AS and PM. The most likely reason for this deterioration in performance is that actual clinical data was used in this study. In practice this means that the recordings contained noise such as friction rubs, rumbling sounds from the stomach, breathing sounds from the lungs and background noise from the clinical environment. In many other studies, the data is either provided from teaching tapes,[2,26] or from specially selected heart cycles of very high quality and with typical morphology.[3,26,41] Another reason could be the rather straight forward choice of parameters used when designing the feed-forward network, or, for that matter, the choice of the classifier. The choice of an optimal classifier was however not the aim of this paper.

The number of patients in the three groups was rather uneven, but this distribution reflects the actual patient population in a Swedish hospital-based echocardiographic laboratory (considering patients in the range middle-aged to elderly). A need for more data is however evident for clinical validation. More patients are also needed since a rule of thumb is to use 10 times as many cases as there are connections in the neural network, which is far from reached in the present set-up.

A weakness with the present system when considering implementation into the intelligent stethoscope is that an ECG is necessary for HS segmentation. Automatic segmentation of the HS signal was tried out on the material without success. When using Shannon energy,[16] the influence of noise and heart murmurs obscured the output signal, even when using wavelet denoising.[17] Automatic segmentation of noisy HS signals still seems to be an open question, and is hereby left for future studies. Another implementation issue is that some of the feature extraction methods are non-causal, which prevents real time implementations. However, all necessary results can be presented to the physician within consultation time using a common computer system.

In conclusion, the results clearly motivate the use of nonlinear features for HS classification. The derived multi-domain SFFS feature subset gives excellent results when compared to single-domain feature sets, and seems rather robust to the noisy data used in this study. It is very interesting that the SFFS subset combines good features from several individual components. For instance, VFD by itself is very good at detecting AS, while it is very poor at detecting MI and PM. ST map and RQA have the best performance detecting PM. The SFFS subset seems to combine good features of all methods presented. Future work includes the design of an optimal classifier and careful clinical tests.

ACKNOWLEDGMENTS

The authors are grateful to the kind and forthcoming personnel at the County Hospital Ryhov in Jönköping, Sweden and at the University Hospital in Örebro, Sweden. This study was supported by the Swedish Agency for Innovation Systems, the Health Research Council in the South-East of Sweden, the Swedish Research Council and the Swedish Heart Lung Foundation.

REFERENCES

[1]Cao, L. Y. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D*. 110:43–50, 1997.

[2]Cathers, I. Neural network assisted cardiac auscultation. *Artif. Intell. Med*. 7:53–66, 1995.

[3]DeGroff, C. G., S. Bhatikar, J. Hertzberg, R. Shandas, L. Valdes-Cruz, and R. L. Mahajan. Artificial neural network-based method of screening heart murmurs in children. *Circulation* 103:2711–2716, 2001.

[4]Durand, L. G., and P. Pibarot. Digital signal processing of the phonocardiogram: Review of the most recent advancements. *Crit. Rev. Biomed. Eng*. 23:163–219, 1995.

[5]Ergen, B., and Y. Tatar. The analysis of heart sounds based on linear and high order statistical methods. *Proc. of the 23rd Ann. Int. Conf. of the IEEE, EMBS*. 3:2139–2141, 2001.

[6]Gao, J. B. Recurrence time statistics for chaotic systems and their applications. *Phys. Rev. Lett*. 83:3178–3181, 1999.

[7]Gnitecki, J., and Z. Moussavi. Variance fractal dimension trajectory as a tool for heart sound localization in lung sounds recordings. *Proc. 25th Ann. Int. Conf. of the IEEE, EMBS* 3:2420–2423, 2003.

[8]Gupta, C. N., R. Palaniappan, S. Swaminathan, and S. M. Krishnan. Neural network classification of homomorphic segmented heart sounds. *Appl. Soft Comp*. In Press, 2006.

[9]Hadjileontiadis, L. J., and S. M. Panas. Discrimination of heart sounds using higher-order statistics. *Proc. 19th Ann. Int. Conf. of the IEEE, EMBS*. 3:1138–1141, 1997.
[10]Hassanpour, H., M. Mesbah, and B. Boashash. Time-frequency feature extraction of newborn EEG seizure using SVD-based techniques. *EURASIP. J. Appl. Sci. Pr.* 2004:2544–2554, 2004.
[11]Kantz, H. and T. Schreiber. Nonlinear Time Series Analysis. Cambridge Univ. Press, Cambridge 369Kantz H., Schreiber T. 2004. Nonlinear Time Series Analysis Cambridge: Cambridge Univ. Press, 369 pp.
[12]Kinsner, W. Tech. Report: Batch and real-time computation of a fractal dimension based on variance of a time series, Univ. Manitoba, Winnipeg, MB, Canada. DEL94-6:1–22, 1994.
[13]Kinsner, W., C. Vincent, K. Cannons, J. Pear, and T. Martin. Signal classification through multifractal analysis and complex domain neural networks. *IEEE Trans. Syst., Man Cybernetics, Part C*. 36:196–203, 2006.
[14]Leung, T. S., P. R. White, W. B. Collis, E. Brown, and A. P. Salmon. Analysing paediatric heart murmurs with discriminant analysis. *Proc. 20th Ann. Int. Conf. IEEE, EMBS*. 3:1628–1631, 1998.
[15]Liang, H. and I. Hartimo. A heart sound feature extraction algorithm based on wavelet decomposition and reconstruction. *Proc. 20th Ann. Int. Conf. IEEE, EMBS*. 3:1539–1542, 1998.
[16]Liang, H., S. Lukkarinen, and I. Hartimo. Heart sound segmentation algorithm based on heart sound envelogram. *Comp. Card.*:105–108, 1997.
[17]Liang, H., L. Sakari, and H Iiro. A heart sound segmentation algorithm using wavelet decomposition and reconstruction. *Proc. 19th Ann. Int. Conf. IEEE, EMBS*. 4:1630–1633, 1997.
[18]Livanos, G., N. Ranganathan, and J. Jiang. Heart sound analysis using the S transform. *Comp. Card.*:587–590, 2000.
[19]Marwan, N., N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths. Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Phys. Rev. E*. 66:1–8, 2002.
[20]Merkwirth, C., U. Parlitz, and L. W. Tstool. A software package for nonlinear time series analysis. *Proc. Int. Workshop Adv. Black-Box Tech. for Nonl. Mod.*, 1998.
[21]Minfen, S. and S. Lisha. The analysis and classification of phonocardiogram based on higher-order spectra. *Proc. of the IEEE-SP, Higher-Order Statistics*.29–33, 1997.
[22]Mohamed, A. S. A. and H. M. Raafat. Recognition of heart sounds and murmurs for cardiac diagnosis. *Proc. 9th Int. Conf. Pattern Recogn*. 2:1009–1011, 1988.
[23]Nigam, V. and R. Priemer. Accessing heart dynamics to estimate durations of heart sounds. *Physiol. Meas.* 26:1005–1018, 2005.
[24]Olmez, T. and Z. Dokur. Classification of heart sounds using an artificial neural network. *Pattern Recogn. Lett.* 24:617–629, 2003.
[25]Pan, J. and W. J. Tompkins. A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng*. 32:230–236, 1985.
[26]Pavlopoulos, S. A., A. C. Stasis, and E. N. Loukis. A decision tree-based method for the differential diagnosis of aortic stenosis from mitral regurgitation using heart sounds. *Biomed. Eng. Online*. 3:21.
[27]Popov, B., G. Sierra, L. G. Durand, J. Xu, P. Pibarot, R. Agarwal, and V Lanzo. Automated extraction of aortic and pulmonary components of the second heart sound for the estimation of pulmonary artery pressure. *Proc. 26th Ann. Int. Conf. IEEE, EMBS*. 1:921–924, 2004.
[28]Povinelli, R. J., M. T. Johnson, A. C. Lindgren, and J. J. Ye. Time series classfication using Gaussian mixture models of reconstructed phase spaces. *IEEE Trans. Knowl. Data En*. 16:779–783, 2004.
[29]Pudil, P., J. Novovicova, and J. Kittler. Floating search methods in feature-selection. *Patt. Recogn. Lett*. 15:1119–1125, 1994.
[30]Quinones, M. A., C. M. Otto, M. Stoddard, A. Waggoner, and W. A. Zoghbi. Recommendations for quantification of Doppler echocardiography: A report from the Doppler Quantification Task Force of the Nomenclature and Standards Committee of the American Society of Echocardiography. *J. Am. Soc. Echocardiogr*. 15:167–184, 2002.
[31]Rajan, S., R. Doraiswami, R. Stevenson, and R. Watrous. Wavelet based bank of correlators approach for phonocardiogram signal classification. *Proc. of the IEEE-SP, Time-Frequency and Time-Scale Analysis.*:77–80, 1998.
[32]Sharif, Z., M. S. Zainal, A. Z. Sha'ameri, and S. H. S. Salleh. Analysis and classification of heart sounds and murmurs based on the instantaneous energy and frequency estimations. *Proc. TENCON* 2:130–134, 2000.
[33]Shino, H., H. Yoshida, K. Yana, K. Harada, J. Sudoh, and E. Harasewa. Detection and classification of systolic murmur for phonocardiogram screening. *Proc. 18th Ann. Int. Conf. IEEE, EMBS* 1:123–124, 1996.
[34]Stockwell, R. G., L. Mansinha, and R. P. Lowe. Localization of the complex spectrum: The S transform. *IEEE Trans. Sig. Proc*. 44:998–1001, 1996.
[35]Taikang, N. and H. Kai-Sheng. Delineation of systolic murmurs by autoregressive modelling. *Proc. of the IEEE 21st Ann. NE. Bioeng. Conf.*:19–21, 1995.
[36]Theodoridis, S. and K. Koutroumbas. Pattern Recognition. Academic Press, Amsterdam 689Theodoridis S., Koutroumbas K. 2003. Pattern Recognition Amsterdam: Academic Press, 689 pp.
[37]Tilkian, A. G. and M. B. Conover. Understanding Heart Sounds and Murmurs: With an Introduction to Lung Sounds. Saunders, Philadelphia 358Tilkian A. G., Conover M. B. 2001. Understanding Heart Sounds and Murmurs: With an Introduction to Lung Sounds Philadelphia: Saunders, 358 pp.
[38]Turkoglu, I., A. Arslan, and E. Ilkay. An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks. *Comput. Biol. Med*. 33:319–331, 2003.
[39]Varanini, M., G. De Paolis, M. Emdin, A. Macerata, S. Pola, M. Cipriani, and C. Marchesi. Spectral analysis of cardiovascular time series by the S-transform. *Comp. Card.*:383–386, 1997.
[40]Webber, C. L. and J. P. Zbilut. Dynamical assessment of physiological systems and states using recurrence plot strategies. *J. Appl. Physiol*. 76:965–973, 1994.
[41]Voss, A., A. Mix, and T. Hubner. Diagnosing aortic valve stenosis by parameter extraction of heart sound signals. *Ann. Biomed. Eng*. 33:1167–1174, 2005.
[42]Xiang, Y. and S. K. Tso. Detection and classification of flaws in concrete structure using bispectra and neural networks. *Ndt. E. Int*. 35:19–27, 2002.