Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Cluster analysis and classification of heart sounds

Guy Amit [a,*], Noam Gavriely [b], Nathan Intrator [a]

[a] *School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel*
[b] *Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel*

ARTICLE INFO

Article history:
Received 6 March 2008
Received in revised form 23 June 2008
Accepted 28 July 2008
Available online 17 October 2008

Keywords:
Heart sounds
Phonocardiography
Cluster analysis
Classification
Time–frequency analysis
Cardiac monitoring

ABSTRACT

Acoustic heart signals, generated by the mechanical processes of the cardiac cycle, carry significant information about the underlying functioning of the cardiovascular system. We describe a computational analysis framework for identifying distinct morphologies of heart sounds and classifying them into physiological states. The analysis framework is based on hierarchical clustering, compact data representation in the feature space of cluster distances and a classification algorithm. We applied the proposed framework on two heart sound datasets, acquired during controlled alternations of the physiological conditions, and analyzed the morphological changes induced to the first heart sound (S1), and the ability to predict physiological variables from the morphology of S1. On the first dataset of 12 subjects, acquired while modulating the respiratory pressure, the algorithm achieved an average accuracy of 82 ± 7% in classifying the level of breathing resistance, and was able to estimate the instantaneous breathing pressure with an average error of 19 ± 6%. A strong correlation of 0.92 was obtained between the estimated and the actual breathing efforts. On the second dataset of 11 subjects, acquired during pharmacological stress tests, the average accuracy in classifying the stress stage was 86 ± 7%. The effects of the chosen raw signal representation, distance metrics and classification algorithm on the performance were studied on both real and simulated data. The results suggest that quantitative heart sound analysis may provide a new non-invasive technique for continuous cardiac monitoring and improved detection of mechanical dysfunctions caused by cardiovascular and cardiopulmonary diseases.

1. Introduction

The activity of the cardiovascular system is periodic by nature. However, as the complex physiological processes driving this system are inherently variable, periodic cardiovascular signals exhibit a considerable beat-to-beat variation. Heart sounds are a good example of periodic, yet variable physiological signals. Changes in the mechanical processes generating the acoustic vibrations, as well as in the propagation medium, cause morphological variations of the acquired signals. In this study we show that an accurate analysis of the signal's morphological variation, using pattern recognition algorithms, can reveal meaningful changes in the underlying physiological processes.

1.1. Heart sounds

The systolic contraction of the ventricles triggers vibrations of the cardiohemic system, including the heart chambers, heart valves and blood. These vibrations propagate through the thoracic cavity and are received on the chest wall as a transient low-frequency acoustic signal, commonly known as the first heart sound, S1. At the end of systole, following closure of the semilunar valves between the ventricles and the arteries, the second heart sound, S2, is produced [1]. The mechanical cardiac cycle is continuously controlled and regulated by the autonomous nervous system, which induces changes to both rate and intensity of myocardial contraction. The pulmonary system has an important part in modulating the cardiovascular activity by respiratory-induced changes in the pleural pressure, pulmonary artery pressure and venous return [2]. The physiological variability of the mechanical function of the heart is reflected in the produced acoustic vibrations—the heart sounds. Heart sounds have been widely used in clinical practice since the introduction of the first stethoscope by Laennec in 1816, and the invention of phonocardiography, the graphic recording of heart sounds, by Einthoven in 1894. Heart sounds and their clinical utilization in cardiovascular and cardiopulmonary diseases have been extensively studied for many years [3]. Relations between morphological features of heart sounds and hemodynamic parameters have been quantitatively described in both animal models and humans [4,5]. Despite the

[*] Corresponding author.
E-mail address: gamit@tau.ac.il (G. Amit).

1746-8094/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bspc.2008.07.003

retained importance of cardiac auscultation in clinical diagnosis [6,7], the use of heart sounds remains mostly qualitative and manual. The immense technological developments of the last decades made cardiac imaging techniques such as echocardiography, computerized tomography (CT) and magnetic resonance imaging (MRI) the state-of-the-art tools of cardiac diagnosis. As much as these imaging technologies are valuable, they require complex equipment and expert operators, and thus cannot be used continuously or outside of the hospital environment. Electrocardiography maintains its central role in cardiac diagnosis and monitoring. The electrocardiogram (ECG) signal provides reliable indications for electrical dysfunctions related to the heart's pacing and conduction systems, as well as for conditions of myocardial ischemia [8]. In particular, ECG-based techniques of heart-rate variability (HRV) have been shown useful in predicting mortality rates in high-risk cardiac patients [9]. However, mechanical dysfunctions that are not accompanied by electrical changes may not be reflected in the electrocardiogram. In addition, patients with chronic heart disease such as heart failure often have enduring ECG abnormalities [10], which reduce the efficacy of ECG monitoring in detecting worsening of the disease. Consequently, long-term non-invasive monitoring of mechanical cardiac function remains unavailable in the common medical practice. In our view, revisiting heart sound analysis using modern computational tools may provide new insights about the relationship between these signals and the mechnical function of the heart and can contribute to improved diagnosis of cardiac malfunctions.

1.2. Analysis techniques

Heart sounds have been previously studied using a variety of digital signal processing techniques, including spectral analysis, parametric and non-parametric time–frequency decomposition and acoustic modeling [11,12]. Signal representation is a fundamental consideration of any signal analysis framework. A representation that is suited to the characteristics of the signal allows a more reliable extraction of features. Heart sounds are low-frequency, non-stationary multi-component signals. Their concurrent variability in both time and frequency domains makes joint time–frequency analysis a favorable method of decomposition and representation. Time–frequency representations, including short-time Fourier transform, Wigner-Ville distribution, continuous wavelet transform and reduced-interference distributions have been previously applied to heart sound signals [13–15]. These non-parametric methods have been shown useful in characterizing the sub-components of the first and second heart sounds and extracting meaningful spectral features from them, with good performance compared to parametric modeling techniques [16]. We have previously used principal component analysis to extract the pattern of spectral changes of S1, associated with the increased cardiac contractility during stress response [17]. Feature extraction is typically a preceding step for a classification or regression task. Heart sound classification, based on morphological spectral and time–frequency features, has been previously used for assessing the condition of bioprosthetic heart valves [18–20]. Classification algorithms used in these studies included *K*-nearest-neighbor, Gaussian–Bayes and neural networks. The reported high accuracy of 89%–98% motivated the utilization and development of advanced pattern recognition techniques for other applications of heart sound analysis. Cluster analysis is a common unsupervised learning technique for partitioning a dataset into subsets of data elements that are similar according to some distance metric. Previous biomedical applications of cluster analysis focused primarily on imaging modalities such as magnetic resonance imaging (MRI) [21]. When applied on periodic physiological signals, cluster analysis can identify groups of signal cycles with distinct common morphologies, as well as point-out outlier noisy cycles with irregular morphologies. If the dataset is labeled, i.e. each signal cycles is associated with a specific physiological condition, a classifier can be constructed to predict the physiological label from the morphology of the signal. The 'correctness' of the clustering is validated in case signal cycles from different clusters are indeed classified into different physiological classes, thus uncovering the relation between morphology and function.

This paper presents a signal clustering and classification framework for inducing parameters of cardiac function from the morphology of acoustic heart signals. The analysis framework is used to study the changes induced to S1 by alternations of the physiological conditions during resistive respiration and during pharmacological stress. We study the signal representation most appropriate for heart sound classification, and compare between time-domain, frequency domain and various joint time–frequency representations, using both real and simulated data. We also address the choice of distance metrics used for comparing heart sound signals, and evaluate the performance of several types of classifiers.

2. Methods

The signal analysis framework used in this study is illustrated in Fig. 1. Heart sound signals (S1) were first identified and extracted from the acquired data, and then transformed to a raw feature space in the time–frequency plane. Hierarchical clustering was applied to the signals, creating a compact representation of the data in the feature space of cluster distances. In this new feature space, classification or regression algorithms were used to test whether the different signal morphologies represent different physiological states.

2.1. Datasets

Two sets of heart sound data were used in this study. In both datasets, heart sounds were acquired during controlled alternations of the physiological conditions, and the ability of the analysis framework to identify the effects of these alternations was evaluated. The physiological variability was induced in the first dataset (HSPRS) by changing the breathing resistance and in the second dataset (HSDSE) by changing the heart's contractility.

1. HSPRS: Heart sounds with alternating breathing resistance. Data consisted of 12 healthy subjects (age 29 ± 12 years, 8 men). For

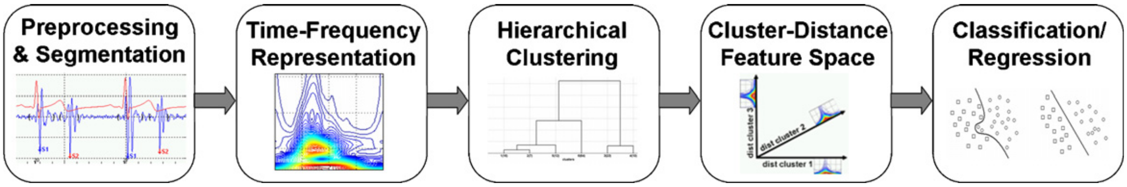
Preprocessing & Segmentation
S1
S2
S1
S2
Time-Frequency Representation
Hierarchical Clustering
Cluster-Distance Feature Space
dist cluster 3
dist cluster 2
dist cluster 1
Classification/ Regression

Fig. 1. Heart sound signal analysis framework.

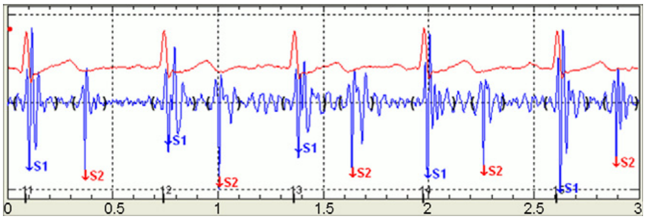

Fig. 2. Segmentation of heart sound signal into cycles of S1 and S2. Heart cycles are determined by the ECG signal. The boundaries of S1 and S2 are marked by black brackets. Consecutive beats of S1 and S2 exhibit noticeable morphological changes.

each subject, two heart sound channels, a single-lead electrocardiogram and breathing pressure were acquired in 10 recordings of 40 s each, while the subject breathed normally against increasing five levels of breathing resistance. Each heart beat was associated with a label of the breathing resistance (0–4) and with the instantaneous measured value of breathing pressure.

2. HSDSE: Heart sounds during pharmacological stress test. Data consisted of 11 male subjects (age 60 ± 14 years) undergoing a routine pharmacological stress test (Dobutamine stress echocardiography). For each subject, four heart sound channels and a single-lead electrocardiogram were continuously recorded during the administration of increasing doses of Dobutamine, which raised the heart's contractility and rhythm. Recording time varied between 30 and 45 min. Each heart beat was labeled by its corresponding stage of the stress test (baseline, 3 to 5 doses of Dobutamine, recovery). A detailed description of the experimental settings and data acquisition can be found in [17].

2.2. Signal representation

The continuous heart sound signal was first pre-processed by applying a digital band pass filter in the frequency range of 20–250 Hz. The signal was then partitioned into cardiac cycles using the peaks of the ECG-QRS complexes as reference points (Fig. 2). The signal segment containing the first heart sound, S1, was defined from 50 ms before the QRS peak to 150 ms after the QRS peak. S1 signals were extracted from each cardiac cycle and aggregated for further processing.

Each beat of S1 was characterized by three types of representation (Fig. 3):

1. Time-domain representation: Direct characterization of the signal as a time series of sampled amplitude values.
2. Frequency-domain representation: Spectral characterization of the signal obtained by applying fast Fourier transform (FFT).
3. Time–frequency representation: Joint time–frequency characterization of the signal obtained by applying one of the following transforms:

- Short-time Fourier transform (STFT), defined by:

$$S(t,f)=\int_{-\infty}^{\infty} s(\tau)W(\tau-t)\,\mathrm{e}^{-\mathrm{i}2\pi f\tau}\,\mathrm{d}\tau$$

- *S*-transform (ST), defined by [22]:

$$S(t,f)=\int_{-\infty}^{\infty} s(t)\frac{|f|}{\sqrt{2\pi}}\,\mathrm{e}^{-(t-\tau)^2 f^2/2}\,\mathrm{e}^{-\mathrm{i}2\pi f\tau}\,\mathrm{d}\tau$$

- Wigner–Ville distribution (WVD), defined by [23]:

$$S(t,f)=\int_{-\infty}^{\infty} s\left(t+\frac{\tau}{2}\right)s^*\left(t-\frac{\tau}{2}\right)\mathrm{e}^{-\mathrm{i}2\pi f\tau}\,\mathrm{d}\tau$$

- Choi–Williams distribution (CWD), defined by [23]:

$$S(t,f)=\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}2\pi ft}\int_{-\infty}^{\infty}\sqrt{\frac{\sigma}{4\pi\tau^2}}\,\mathrm{e}^{-\sigma(\mu-t)^2/4\tau^2}\times s\left(\mu+\frac{\tau}{2}\right)s^*\left(\mu-\frac{\tau}{2}\right)\mathrm{d}\mu\,\mathrm{d}\tau$$

where $s(t)$ is the original signal, t the time delay, f the frequency, W a window function and σ is a parameter controlling the suppression of cross-terms.

2.3. Cluster analysis

Hierarchical clustering was applied to S1 signals, using each of the signal representations described above. The purpose of clustering is to partition a dataset into disjoint subsets (clusters), such that data elements within the same cluster share some sort of similarity. Similarity between data elements is measured using a distance metric that is suitable for the nature of the analyzed data. Two distance metrics were considered in this study:

1. Euclidean distance: $D_{sr}=||s_t-r_t||^2=\sum_t(s_t-r_t)^2$, where s_t, r_t are signals of length n.
2. Cross-correlation:

$$D_{sr}=1-\frac{\sum_t(s_t-\bar{s})(r_t-\bar{r})}{\sqrt{\sum_t(s_t-\bar{s})^2}\sqrt{\sum_t(r_t-\bar{r})^2}}$$

where $\bar{s}=1/n\sum_{t=1}^{n}s_t$, $\bar{r}=1/n\sum_{t=1}^{n}r_t$

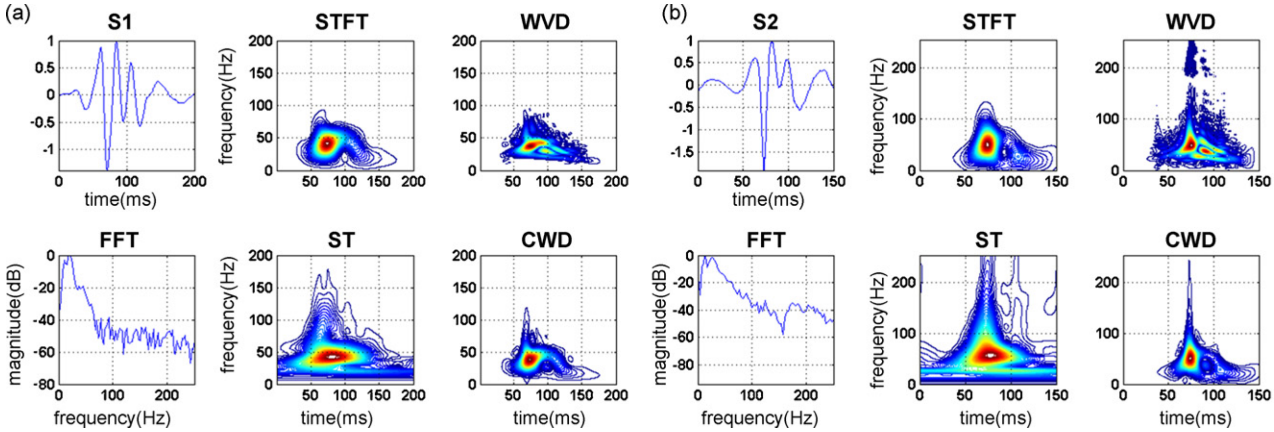

Fig. 3. Representation of S1 (a) and S2 (b) in the time-domain, frequency-domain (FFT) and by various joint time–frequency transforms: short-time Fourier transform (STFT), *S*-transform (ST), Wigner–Ville distribution (WVD) and Choi–Williams distribution (CWD). STFT has fixed resolution, while ST has frequency-dependent resolution. WVD has higher resolution, but its quadratic nature creates cross-terms, which are suppressed in the reduced-interference CWD representation.

Clustering was done using an agglomerative hierarchical clustering procedure [24] that initially partitions a set of n data elements into n clusters, each containing one data element, and then iteratively merges the two most similar clusters, until the entire dataset forms a single cluster. The bottom of the created hierarchical tree can next be pruned so that the required number of clusters N is obtained. Data elements below each cut are assigned to a single cluster, creating the output data partitioning to clusters $\{C_1, \ldots, C_N\}$. The algorithm requires a cluster similarity criterion for choosing the next two clusters to be merged. We have used Ward's step-wise optimal criterion, which chooses the clusters such that the increase in the overall sum-of-squared error after the merge is minimal [25]. The distance between clusters C_i and C_j is defined by: $D_W(C_i, C_j) = \sqrt{n_i n_j / n_i + n_j} || m_i - m_j ||$, where n_i, n_j are the sizes of clusters, and m_i, m_j are their means.

2.4. Classification framework

While cluster analysis is used to identify distinct signal morphologies in the data, the classification framework is aimed to uncover the relation between these morphologies and the alternating physiological conditions. This is achieved by evaluating the ability of a classifier to accurately predict the label of each heart beat, given only its morphological representation. The input of the clustering-classification framework is a dataset of n heart sound cycles, $B = \{(b_1,l_1),(b_2,l_2), \ldots, (b_n,l_n)\}$, where b_i is a representation of a heart sound component (e.g. S1) during a single cardiac cycle, and l_i is its associated class label $l_i \in \{L_1, \ldots, L_m\}$. The cluster analysis procedure assigns a cluster identifier to each signal cycle, producing a clustered dataset $C = \{(b_1,c_1), (b_2,c_2), \ldots, (b_n,c_n)\}$, where $c_i \in \{1, \ldots, N\}$ are arbitrary cluster identifiers. Using these notations, a cluster C_j is the set of signal cycles with cluster identifier c_j: $C_j = \{i|(b_i,c_j) \in C\}$. Clusters that contain a minimal portion of the data, i.e. $|C_j| \geq \beta n$, are denoted as *significant* clusters. β was set by experiment to 0.05. The center of a cluster C_j is a weighted average of the cluster's elements, in which each signal cycle is weighted by its similarity to the cluster's arithmetic mean: $\bar{C}_j = \sum_{i \in C_j} \omega_i b_i$, $\omega_i = 1 - D(b_i, (\sum_{i \in C_j} b_i)/|C_j|)$, where D is a distance function.

The centers of the significant clusters provide a compact representation of the morphological variability in the entire dataset. Furthermore, a signal cycle b_i can be efficiently characterized by the vector of its distances from the centers of the significant clusters $\vec{d}^i = (d_1^i, d_2^i, \ldots, d_{\hat{N}}^i)$, $d_k^i = D(b_i, \bar{C}_k)$. The classification algorithm is applied in this new feature-space of cluster distances.

The outline of the classification framework is as follows:

1. Classification is applied separately on the dataset B of each subject.
2. Data is partitioned into a training set B^{train} and a testing set B^{test}.
3. Hierarchical clustering is applied on the training set, producing clustered data C^{train}.
4. The centers of the significant training clusters $\bar{C}_1^{\text{train}}, \ldots, \bar{C}_{\hat{N}}^{\text{train}}$ are calculated.
5. Each beat $b_i \in B^{\text{test}} \cup B^{\text{train}}$ is characterized in the cluster-distance space by the vector $\vec{d}^i = (d_1^i, d_2^i, \ldots, d_{\hat{N}}^i)$ of its distances from the centers of the significant training clusters.
6. A classifier F is constructed from the cluster distance-space representation of the training set. For beat $b_i \in B^{\text{train}}$, $F(b_i) = F(d_1^i, d_2^i, \ldots, d_{\hat{N}}^i) = \tilde{l}_i$, $\tilde{l}_i \in \{L_1, \ldots, L_m\}$.
7. The classification accuracy is evaluated on cluster distance-space representation of the testing set B^{test}.

Two classification methods were used, and their performances were compared:

1. K-nearest-neighbor (KNN): A non-parametric method that classifies a test data element by a majority vote of the closest data elements in the training set [26]. Given a labeled training set B and a test data element d, KNN classifies d by choosing K train data elements $\{b_1, \ldots, b_K\} \subset B$ that are the closest neighbors of d under a distance metric D: $D(d,b_1) \leq D(d,b_2) \leq \ldots \leq D(d,b_K) \leq D(d,b_j)$. $\forall b_j \in Bj \notin \{1, \ldots, K\}$. Then, given that l_i is the label of train data element b_i, d will be classified as the statistical mode of $\{l_1, \ldots, l_K\}$. In the regression case, when the training data elements are associated with real values ν_i rather than discrete class labels, the value associated with d is estimated as a weighted average: $v(d) = \sum_{i=1}^K D(d, b_i)\nu_i / \sum_{i=1}^K D(d, b_i)$.
2. Discriminant analysis (DA): Finds a linear transform that maximizes the separation between classes in the training set [27]. The maximized objective function is: $J(w) = w^t S_B w / w^t S_W w$, where $S_B = \sum_{c=1}^m n_c(\bar{d}_c - \bar{d})(\bar{d}_c - \bar{d})^{\mathrm{T}}$, $S_W = \sum_{c=1}^m \sum_{i \in c}(d^i - \bar{d}_c)(d^i - \bar{d}_c)^{\mathrm{T}}$ are the between-classes scatter matrix and the within-class scatter matrix, respectively, n_c is the number of data elements in class c, $\bar{d}_c = 1/n_c \sum_{i \in c} d^i$ is the mean of class c and $\bar{d} = (1/N)\sum_i d^i = (1/N)\sum_{c=1}^m n_c \bar{d}_c$ is the mean of the entire training set. Once the transformation w is found, by solving an eigenvalue problem, a test data element d can be classified to $\arg\min_c D(dw, \bar{d}_c w)$, the class whose center is closest to d, under a distance metric D.

The distance metric used by the classification algorithm is the mahalanobis distance, defined by $D(\vec{d}^i, \vec{d}^j) = (\vec{d}^i - \vec{d}^j)V^{-1}(\vec{d}^i - \vec{d}^j)^{\mathrm{T}}$, where V is the covariance matrix of vectors $\vec{d}^i$ and $\vec{d}^j$.

Since the class labels in this data represent a continuum of physiological changes, rather than dichotomic classes, the classification accuracy CC_m was defined as the percentage of data elements classified within a certain range m of their actual label (typically, $m = 1$):

$$\mathrm{CC}_m = \frac{|\{b_i \in B^{\text{test}} || \tilde{l}_i - l_i| \leq m\}|}{B^{\text{test}}}$$

2.5. Classification of heart sound data

The classification framework was applied separately on the records of each subject from the HSPRS and HSDSE datasets. For each of the five levels of breathing resistance in the HSPRS dataset, one 40-s signal was used for training and one 40-s signal was used for testing. Analysis was carried out on S1 signals from a single heart sound channel. Each beat was associated with the label of its corresponding breathing resistance (five classes: R_0–R_5), as well as with a real value of the instantaneous breathing pressure. To compute the instantaneous breathing pressure, cluster analysis was performed separately on the data of each resistance level, producing significant clusters $C_1, \ldots, C_{\hat{N}}$ that are related to the phase of the respiratory cycle (inspiration and expiration). The value of the breathing pressure associated with a signal cycle i was defined as $\nu_i = P(t_i - \tau)$, where t_i is the reference time point of signal cycle i (typically, the beginning of the cycle), and τ is a constant delay parameter ($0 \leq \tau \leq 800$ ms). The delay parameter τ was chosen to provide the maximal separation between the pressure values of the significant clusters, in terms of Fisher's separation criterion FC [27]:

$$\mathrm{FC} = \frac{\sum_{j=1}^K p_j(m_j - \bar{m})^2}{\sum_{j=1}^K p_j S_j}, \text{ where } m_j = \frac{1}{|C_j|}\sum_{i \in C_j} \nu_i$$

$$S_j = \frac{1}{|C_j| - 1}\sum_{i \in C_j} (\nu_i - m_j)^2, \; p_j = \frac{|C_j|}{\sum_k |C_k|}, \; \bar{m} = \frac{1}{\sum_k |C_k|} \sum_{i \in \{C_1, \ldots, C_{\hat{N}}\}} \nu_i$$

For each stage of the stress test in the HSDSE dataset, a consecutive 1/3 of the data was used for testing, and the remaining 2/3 of the data was used for training. Consecutive beats were selected in order to ensure unbiased representation of beats occurring at different phases of the respiratory cycle, since there are respiratory-induced morphological variations of S1.

Each beat was associated with the label of the corresponding stress-test stage. The number of classes varied from 5 to 7 between subjects. All subjects had a 'baseline' and 'recovery' classes, and a varying number of stress stages. Analysis was performed separately on S1 signals from each of the four heart sound channels. Classification results from all four channels were combined by a majority vote scheme. Noisy test beats that were assigned to non-significant clusters in more than two channels were excluded (mean 31 ± 13 test beats per subject, 3.7% of the test set). Cluster analysis was applied on the training data with the required number of clusters set to 16. Before clustering, the signals were aligned by shifting each cycle to maximize the cross-correlation with an arbitrary reference cycle. Significant clusters were defined as clusters containing at least 5% of the data. Label classification was done using either KNN with $K = 5$ and mahalanobis distance or DA with mahalanobis distance. Classification performance was evaluated by computing CC_1. Pressure estimation on the HSPRS data was done using KNN. The estimated pressure was the weighted average of the nearest neighbors. The mean pressure estimation error, relative to the peak-to-peak amplitude of pressure variation was computed by: $\mathrm{EE} = (1/|B^{\text{test}}|)\sum_{b_i \in B^{\text{test}}} |\tilde{p}(i) - p(i)|/A_{l_i}$, where $p(i)$ and $\tilde{p}(i)$ are the reference and the estimated instantaneous pressure values of beat i, and A_{l_i} is the peak-to-peak variation of pressure amplitude when breathing against resistance $l_i \in \{R_1,R_2,R_3,R_4\}$. The linear correlation between the instantaneous pressures $\tilde{p}(i)$ and $p(i)$ was also calculated, as well as the correlation between the peak-to-peak breathing amplitude A_{l_i} and its estimation $\tilde{A}_r$, defined by: $\tilde{A}_r = \max\{\tilde{p}(i)|l_i = r\} - \min\{\tilde{p}(i)|l_i = r\}$, $r \in \{R_1,R_2,R_3,R_4\}$.

3. Results

3.1. Heart sounds and breathing pressure

The average number of heart beats processed per subject in the HSPRS dataset was 524 ± 75. Cluster analysis performed separately on data of each breathing resistance level was able to identify distinct morphologies of S1, regardless of the chosen signal representation and distance metric. Signal averaging within the clusters exhibited small morphological variability, compared to averaging of the unclustered data, and thus provided more accurate description of the data (Fig. 4). A strong association was observed between the identified clusters and the respiratory phase: heart beats that followed a high breathing pressure during expiration were morphologically different than beats that followed a low breathing pressure during inspiration (Fig. 5). When applied on the entire data of each subject, the clustering procedure produced a compact feature space of cluster-distances. In this feature space, the separation between heart beats associated with different respiratory phases and different breathing resistance levels could be clearly observed (Fig. 6). The classification algorithm was then used to quantitatively assess the relation between the morphology of the sound signal and the respiratory pressure. The number of significant clusters identified by the analysis framework varied from 6 to 13 clusters per subject (mean 10 ± 2). The classification performance was similar for beats recorded at all breathing resistance levels. For S1 signals, the average correct classification rate CC_1 varied from 72% to 82%, and the pressure estimation error EE varied from 19% to 23% (Table 1). The performance differences between the methods were not very large: correlation distance was somewhat better than Euclidean distance, and KNN classifier was slightly better than DA. The best correct classification rates (82 ± 7%) were obtained by time-domain and *S*-transform representations. All representation methods achieved low estimation errors, with the best result of 19 ± 6% achieved by WVD. A good correlation was obtained between the breathing pressure estimated from the morphology of S1 and the pressure value associated with each beat. The correlation coefficient was 0.76 for the 2057 test beats of all 12 subjects (Fig. 7a). In addition to instantaneous pressure estimation, the peak-to-peak amplitude of the estimated pressure in each breathing resistance level was strongly correlated with the actual pressure variation, or the actual breathing effort ($R = 0.92$, Fig. 7b). To ascertain that these relations are indeed a consequence of the morphological differences between beats, correctly derived by the classification framework, the selection of the *K* nearest neighbors in the cluster-distance space was replaced by a random selection of *K* training beats that were used for classification and pressure estimation. Using this random classification, the results were significantly worse ($p < 10^{-5}$) with average CC_1 of 52%, and average EE of 35%. There was no significant correlation whatsoever between the randomly estimated and the actual pressure per beat.

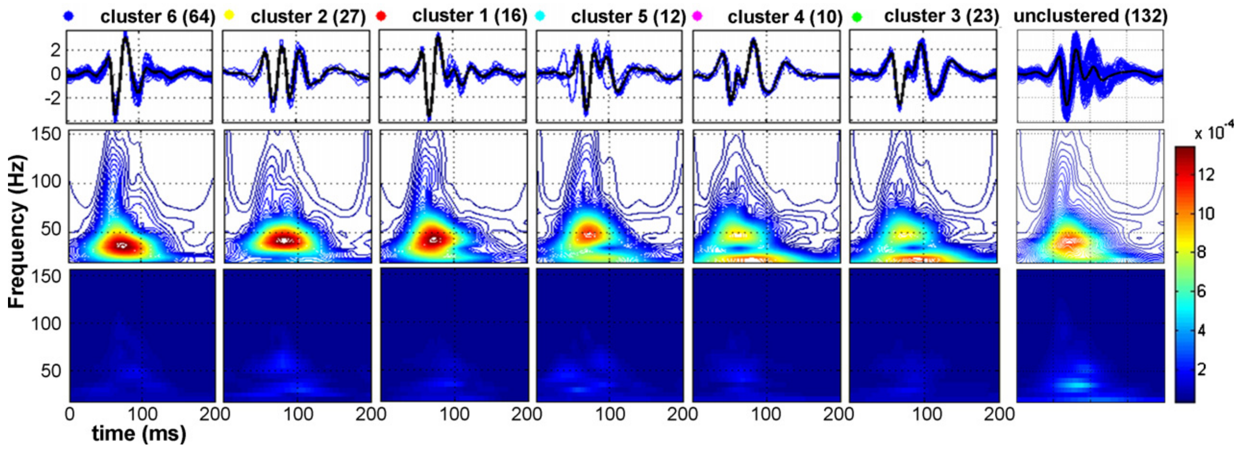

Fig. 4. Clustering results of 132 beats of S1, recorded during strenuous breathing against high resistance. The distinct average signal morphologies and the low standard deviation of the clustered signals are apparent both in the time-domain plot (top) and in the *S*-transform representations (middle and bottom), compared to the unclustered data (right column).

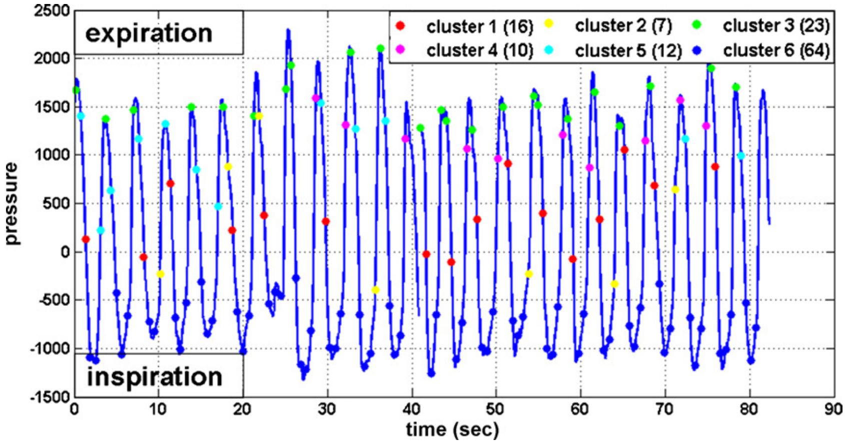

Fig. 5. The relationship between clusters of 132 beats of S1 and the instantaneous breathing pressure. Clusters 3 and 4 are associated with high pressure values (expiration), while cluster 6 is associated with low pressure values (inspiration). The remaining clusters are associated with intermediate pressure values. The phase lag between the occurrence of S1 and the pressure waveform was 455 ms.

3.2. Heart sounds and stress response

In the HSDSE dataset, the average number of processed heart beats per subject was 2549 ± 759. The number of significant clusters identified by the cluster analysis procedure varied from 4 to 10 per subject (mean 7 ± 2). A considerable association was observed between the clusters and stages of the stress test, where each stage was dominated by two to three clusters (Fig. 8). The same clusters were associated with the baseline and the recovery stages, indicating that the observed morphological changes were indeed induced by the stress response. Examining the average morphology of the detected clusters revealed a pattern of increase in the spectral energy and bandwidth, directly related to the stress level. This pattern, obtained by an unsupervised learning technique, is consistent with our previous findings about stress-induced changes of S1 [17]. Representation of beats in the feature space of cluster distances provided a good separation between beats from different test stages (Fig. 9). The observed change in the cluster-distance representation of the S1 was gradual and smooth, becoming more profound at higher stress stages, and returning back to the baseline morphology at the later stage of recovery. The average rates of correct classification (CC_1) of S1, achieved by different combinations of signal representations, distance metrics and classifiers on all 11 subjects varied from 77% to 86% (Table 2). Correlation distance performed better than Euclidean distance, and DA classifier was slightly superior to KNN classifier. The best average classification performance of 86 ± 7% was achieved by the DA classifier on signals represented by the *S*-transform and clustered using correlation distance. Time-domain representation, with correlation distance and DA classifier, provided equivalently good performance, with correct classification of 85 ± 8%. Frequency-domain representation was inferior, compared to time-domain or joint time–frequency representations. No significant differences were observed between STFT, WVD and CWD.

3.3. Simulated signals

In order to realize the differences between signal representations, a simple simulation was conducted. The baseline simulated signal was a 30 Hz sinus wave with duration of 300 ms and a Gaussian amplitude modulation. Random noise with Gaussian distribution was added to the signal, with initial signal-to-noise ratio set to 7 dB. Three types of signal transformations were simulated: (i) time shift between −200 and +200 ms, (ii) frequency change between 20 and 40 Hz and (iii) SNR change between −6 and 10 dB. For each of the transformations, the correlation distance between the baseline signal and the transformed signal was

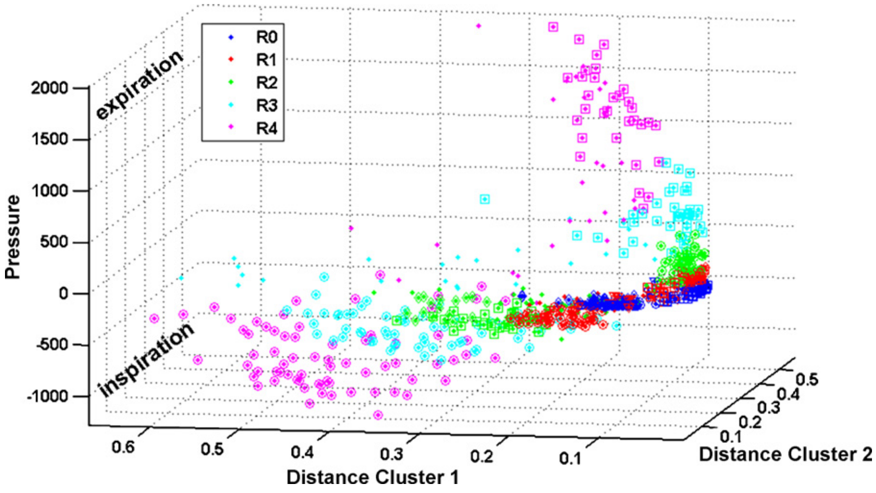

Fig. 6. Cluster-distance representation of 652 beats of S1 from a single subject. Each beat is plotted by its distances from the centers of the two largest clusters (*x*-axis and *y*-axis) and by its associated breathing pressure (*z*-axis). The marker colors show the breathing resistance levels and the marker symbols designate the significant clusters of each level. There is a marked separation between beats of different resistance levels, and within each level, between beats associated with different respiratory phases.

Table 1
Classification performance on S1 signals from HSPRS dataset

Signal representation	Distance metric	KNN CC_1 (%)	DA CC_1 (%)	EE (%)
Time	Correlation	**82 ± 7**	76 ± 11	20 ± 7
	Euclidean	82 ± 7	80 ± 9	21 ± 6
Frequency	Correlation	77 ± 6	72 ± 9	20 ± 4
	Euclidean	75 ± 7	74 ± 5	23 ± 5
Short-time Fourier transform (STFT)	Correlation	78 ± 10	73 ± 10	21 ± 5
	Euclidean	78 ± 8	76 ± 8	23 ± 6
S-transform (ST)	Correlation	**82 ± 7**	76 ± 10	20 ± 7
	Euclidean	78 ± 8	80 ± 9	22 ± 5
Wigner–Ville distribution (WVD)	Correlation	81 ± 7	77 ± 8	**19 ± 6**
	Euclidean	80 ± 7	79 ± 8	20 ± 6
Choi–Williams distribution (CWD)	Correlation	78 ± 8	73 ± 9	20 ± 6
	Euclidean	78 ± 8	76 ± 9	22 ± 7

Mean and standard deviation of correct classification (CC_1) and relative estimation error (EE) of all subjects, using different configurations of signal representation, distance metric and classification algorithm (KNN = *K*-nearest neighbor, DA = discriminant analysis). Best results, obtained by ST, WVD and time-domain representations, are indicated by boldface.

calculated (Fig. 10). Spectral signal representation was obviously insensitive to time shifts. Time representation, on the other hand, was over sensitive, as the distance in this case is the autocorrelation function, which fluctuates between high positive and negative values. Time–frequency representations were more robust to temporal shifts, providing a smooth change of the distance (Fig. 10a). The sensitivity to changes in the signal's frequency was higher for WVD, spectral and time-domain representations, compared to ST and STFT, which have lower frequency resolution (Fig. 10b). Finally, lower signal-to-noise ratio affected ST and WVD much more than STFT and spectral representations, with mediocre noise sensitivity of the time-domain representation (Fig. 10c).

4. Discussion

The relations between the physiological processes producing the heart sounds and the morphology of the externally acquired acoustic signals are highly complex. The mechanical interplay between myocardial contraction, blood flow and valve activity is continuously regulated by the autonomous nervous system, and is affected by hormonal and pulmonary activities. The filtering effects of the thoracic cavity and the skin conducting the acoustic vibrations considerably alter the morphology of the signal [28]. Nevertheless, the heart sounds, being a direct manifestation of the mechanical cardiac cycle, bear valuable information about the functioning of the cardiovascular system. Extracting this information using pattern recognition techniques may provide new means of continuous monitoring and assessment of cardiovascular mechanical function. In order to successfully predict the physiological condition from the morphology of the signal, the analysis techniques should be tailored to fit the properties of the analyzed signals. There is no consensus in the literature regarding the most suitable time–frequency representation of S1. Different studies point out different techniques such as the binomial transform [13], cone-kernel distribution [15] and continuous wavelet transform [14] as the best choices. Indeed, when considering the problem of accurate decomposition of the signal into its subcomponents, there are marked differences between methods. Simple STFT is limited by its fixed resolution, which imposes a tradeoff between temporal and spectral resolutions. One way to avoid the resolution tradeoff is to use linear transforms with frequency-dependent resolution such as the wavelet transform and *S*-transform. Alternatively, quadratic transforms such as WVD and its reduced-interference derivatives like CWD, can be used. In our analysis framework, the preservation of the relative morphological similarity between signals, under a certain representation method, is more important than the absolute accuracy of the signal's decomposition. The optimal signal representation and distance metric should have the correct balance between sensitivity and robustness. Sensitivity is important for detecting minute differences between beats, and robustness is essential to reject noise-related differences. The classification results obtained by simple time-domain representation were comparable in most cases with the results obtained by time–frequency representations (TFR). Theoretically, both types of

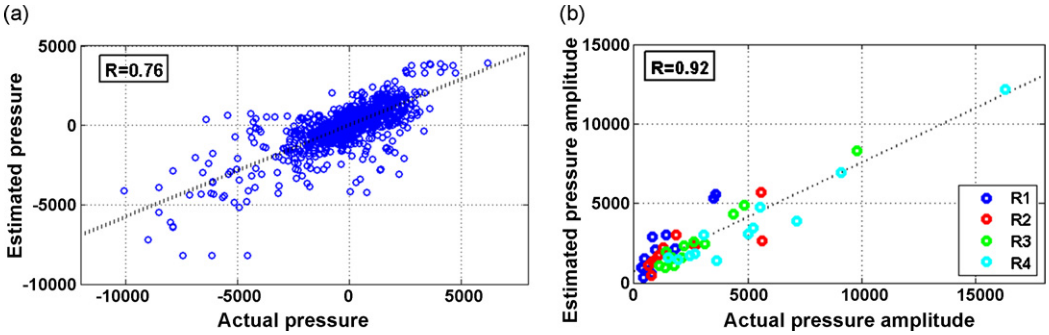

Fig. 7. (a) Estimated breathing pressure of 2057 test beats of S1 from all 12 subjects, plotted against the actual breathing pressure associated with each beat. The correlation coefficient is 0.76. The absolute pressure differences were normalized by the peak-to-peak amplitude of pressure variation to obtain the reported average estimation error. (b) Estimated peak-to-peak amplitude of breathing pressure of 12 subjects (4 resistance levels per subject, indicated by marker symbols), plotted against the measured amplitude of breathing pressure. The correlation coefficient is 0.92. Pressure values are specified in arbitrary non-calibrated transducer units.

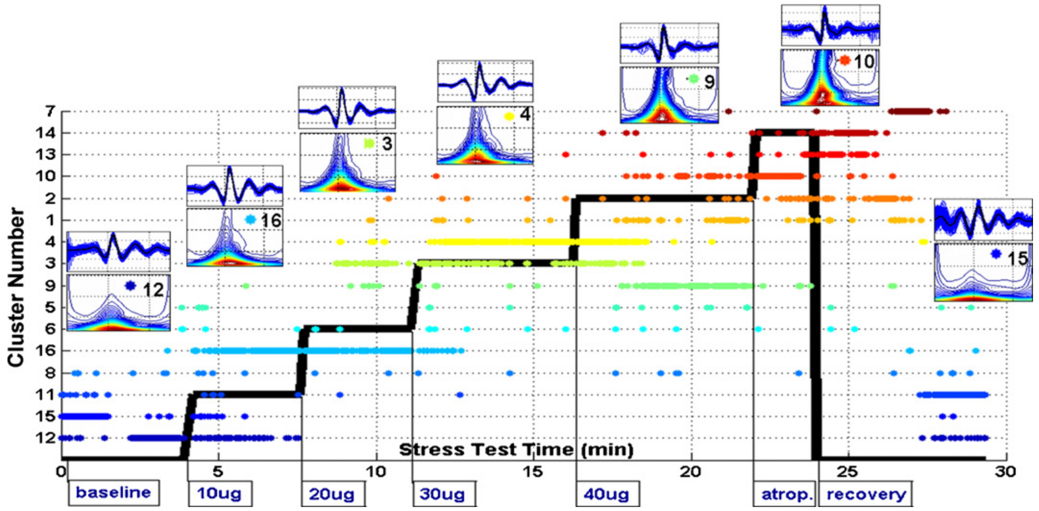

Fig. 8. Clustering results of 2725 beats of S1 acquired from a single subject during 29 min of Dobutamine stress test. Clusters are marked by different colors and by number labels on the *y*-axis. The stress level is represented by the bold black line, labeled with the test stages. The time-domain and *S*-transform representations of the significant clusters exhibit substantial morphological changes, strongly associated with stages of the stress test, with a return to the baseline morphology during recovery.

representations hold the same amount of information about the signal. In TFR, this information is represented in two dimensions, with the cost of either sub-optimal time/frequency resolution, or interference of artifactual cross-terms. When comparing a pair of signals, the 2D representation is more robust to small alignment differences between the signals, and there are significant differences in the sensitivity to changes of the frequency and the noise level (Fig. 10). The choice of signal representation is therefore tightly related to the nature of the variability in the data. In cases where the data exhibits large variability between classes and small variability within each class, highly sensitive representations would provide more accurate results, whereas when the changes in the data are more gradual and there is small between-class variability or large within-class variability, a representation that is less sensitive but more robust should be preferred.

An important aspect of the proposed clustering-classification framework is the compact representation of the data in the feature space defined by the distances from the centers of the significant clusters. It is impractical to use raw signal representation for classification as the data is too high-dimensional. Previous studies on heart sound classification used domain-specific features such as dominant frequencies, spectral bandwidth and signal intensities [18]. More general feature extraction techniques used either model estimation [19] or search-based feature selection [20]. While the domain-specific features have physical meaning and can therefore be easily interpreted, they need to be specifically determined for every type of signal and for every dataset. Automatic feature extraction and selection methods provide a more systematic solution, but their loose relation with the underlying physiological processes makes the classification results less traceable. The weighted averages of the significant clusters constitute a concise description of the most prominent signal morphologies. Cluster belonging alone, as a one-dimensional feature, is too coarse to reliably classify a data element. As there is also considerable within-cluster morphological variability, and since boundaries between similar clusters might be arbitrary, characterizing each data element by measuring its distances from multiple cluster centers is extremely informative, yet computationally efficient. For

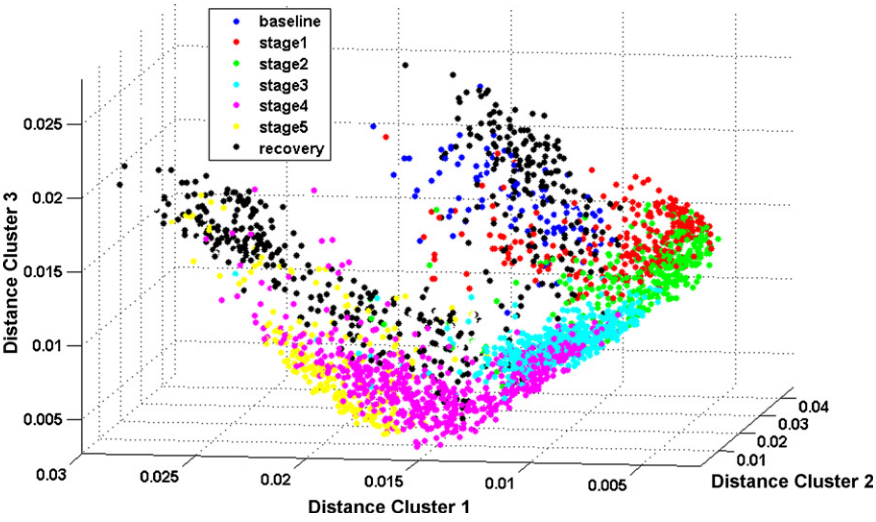

Fig. 9. Cluster-distance representation of 2725 beats of S1 from a single subject, plotted by their distances from the centers of the three largest clusters. The marker colors indicate the stage of the beat in the stress test (baseline, 5 ascending stress levels and recovery). The morphology of S1 seems to vary smoothly along the stages of the test, with a distinct separation between beats of consecutive stages and a return to the baseline morphology towards the end of recovery.

Table 2
Classification performance on S1 signals from HSDSE dataset

Signal representation	Distance metric	KNN CC_1 (%)	DA CC_1 (%)
Time	Correlation	81 ± 8	**85 ± 8**
	Euclidean	81 ± 9	82 ± 8
Frequency	Correlation	77 ± 9	77 ± 11
	Euclidean	75 ± 8	77 ± 9
Short-time Fourier transform (STFT)	Correlation	80 ± 8	80 ± 8
	Euclidean	78 ± 10	80 ± 9
S-transform (ST)	Correlation	**85 ± 7**	**86 ± 7**
	Euclidean	84 ± 0	84 ± 8
Wigner–Ville distribution (WVD)	Correlation	80 ± 9	82 ± 9
	Euclidean	79 ± 9	80 ± 9
Choi–Williams distribution (CWD)	Correlation	80 ± 9	82 ± 9
	Euclidean	79 ± 9	80 ± 9

Mean and standard deviation of correct classification measure (CC_1) of all subjects, using different configurations of signal representation, distance metric and classification algorithm (KNN = *K*-nearest neighbor, DA = discriminant analysis). Best results, obtained by ST and time-domain representations, are indicated by boldface.

heart sounds, typical distance-space representation had 7–10 dimensions, while raw time–frequency representation of S1 had typically 3500 features (100 time points × 35 frequency bins). Distance-space representation is a general approach, applicable to different types of signals, and at the same time it has a simple physical interpretation of morphological signal similarity, which can be partially visualized using three-dimensional plots (Figs. 4 and 5).

Application of the analysis framework on experimentally controlled heart sound data provides some physiological insights about the nature of the morphological variability of heart sounds. Analysis of the HSPRS data demonstrated the effects of the respiratory activity on the morphology of the first heart sound. The phase of the respiration cycle (inspiration or expiration), indicated by the instantaneous breathing pressure, has a marked effect on the heart sound signal. The unsupervised clustering algorithm produced clusters that were separable by their associated breathing pressure levels (Fig. 5). The best separation between high-pressure (expiration) and low-pressure (inspiration) clusters was observed when a delay of 200–600 ms was taken between the measure of the instantaneous pressure and the resulting morphology change. This indicates that the modulation of the sound is a result of a regulation process, which involves the changes in venous return, the volumes of the right and left ventricles and the force of myocardial contraction [2]. The increase of the breathing resistance makes the effects of respiration more prominent. The morphological changes induced by the resistance level are minor compared to the respiration-induced changes. The average estimation error of the breathing pressure from the morphology of S1 was roughly 20%, regardless of breathing resistance. This indicates that there is indeed a strong relation between the breathing pressure and the time–frequency morphology of the first heart sound. This relation may be too complex for an accurate direct modeling, but using computational learning methodology it is possible to describe it and make fairly accurate predictions of the respiratory pressure. This may have implications for non-invasive assessment of cardiopulmonary diseases.

The results obtained on the HSDSE dataset demonstrate the robustness of the proposed framework on data acquired in a realistic clinical environment. The induction of the pharmacological stress agent Dobutamine increases the cardiac contractility, causing significant changes to the first heart sound. We have previously characterized these changes of S1 as an increase in the spectral energy, accompanied by an increase in the frequency bandwidth as higher frequency components in the

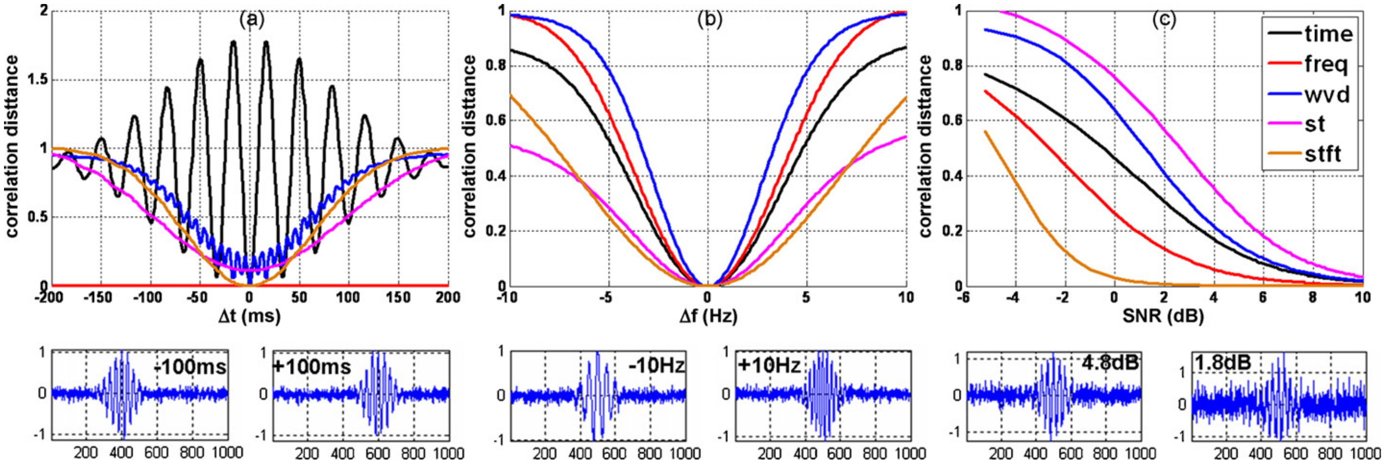

Fig. 10. The sensitivity of the correlation distance under different signal representations to simulative changes of the temporal location (a), the frequency content (b) and the signal-to-noise ratio (c). The baseline simulated signal is a 300 ms, 30 Hz sinus with a Gaussian amplitude modulation and additive Gaussian white noise of 7 dB. Bottom panels present examples of the simulated signals. Time–frequency representations are relatively robust to temporal shifts, compared to over-sensitivity of time-domain representation and insensitivity of spectral representation (a). WVD, spectral and time-domain representations are more sensitive to frequency changes than ST and STFT (b). WVD and ST are more sensitive than STFT to changes in the signal-to-noise ratio (c).

range of 50–150 Hz emerge and strengthen [17]. The current results validate the relation between the Dobutamine dose, which reflects increased cardiac contractility, and the time–frequency morphology of S1. The clustering and classification framework produced an excellent separation between beats at different stages of the stress test (Fig. 9), confirmed by the ability to predict the stage of the test from the signal's morphology with high average accuracy of 86%. The utilization of multiple heart sound channels, acquired simultaneously from different locations, has an important contribution to the high classification accuracy. While each separate channel provided a lower classification rate of about 80%, the fusion of the four classifiers made the analysis more robust and more accurate. This approach can be extended to combining classifiers that use different signal representations, distance metrics or classification algorithms to further improve the accuracy.

On HSPRS dataset, KNN classification provides better results than DA. Opposite results were obtained for HSDSE dataset, where DA was consistently better. This can be explained by the different structures of the cluster spaces created for the two datasets. In the HSDSE data, the Dobutamine-induced stress causes substantial spectral changes of S1 that are reflected as extremely distinct clusters, separable by a linear projection. The respiratory-induced variations in this dataset are masked by the dominant stress-induced changes, and have a negligible influence on the clustering results. The morphological changes of S1 in the HSPRS data are more subtle. Most of the variation is related to the respiratory phase, while the variation caused by changes in the breathing resistance is smaller. Beats of different breathing resistances are therefore intermixed in the cluster-distance space, and the non-linear separation of KNN achieves better results than linear DA.

One of the limitations of this study is the difficulty to distinguish changes in the heart sounds triggered by the modulations of the physiological processes from changes caused by external factors such as body movements and ambient noise. As the signals were externally recorded, they constitute an indirect and possibly distorted representation of the producing processes. This is almost fully handled by the clustering algorithm, which identifies irregular morphologies and assigns them to insignificant clusters. Significant clusters contain only recurrent signal morphologies, and consequently irregular or noisy signals are expressed as outliers in the feature space of cluster distances, and are eliminated from the analysis. Another limitation is that the data representation in the framework in inherently relative. Each beat is characterized by its similarity to the prominent signal morphologies, without any absolute measure of its properties. As there is a large inter-subject variability of the signal's morphology it seems unrealistic to generalize a single classification framework for multiple subjects. This requires training of the framework separately for each subject. The clustering algorithm was not much affected by the choice of the distance metric. The choice of standard distance measures such as Euclidean or correlation distances is a compromise on the accuracy of the signal similarity measure, with the benefit of simplicity and efficiency. These distance measures preserve similarity for simple signal transformations such as amplitude scaling, but they do no account for more complex transformations such as scaling of the time axis. The development of an advanced similarity measure, specifically suited for multi-component, non-stationary signals such as the first heart sound, is a future research direction, expected to further improve the accuracy of the analysis framework. The utilization of ECG for cycle segmentation and temporal location of S1 was another simplifying choice, taken to ensure reliable, straightforward signal segmentation. However, methods for heart sound segmentation without ECG have been proposed [29], and their incorporation in the analysis framework was left for future work.

Monitoring the dynamics of the mechanical cardiac function continuously and non-invasively is an important clinical application that is still unavailable in the common medical practice. The proposed framework for analysis of acoustic heart signals can be straightforwardly applied for patient monitoring. The framework should be first trained on the patient's signals acquired in controlled conditions (for example, during a stress test). Following training, the patient can be continuously monitored by classifying each beat into the appropriate class of physiological condition (e.g. contractility level), or estimating an associated physiological quantity (e.g. pleural pressure). Such application can improve the diagnosis and management of cardiac and respiratory dysfunctions.

5. Conclusions

We have described a signal analysis framework for heart sounds, which consists of time–frequency signal representation, hierarchical clustering, cluster-distance feature space and classification algorithm. The framework, applied on two datasets of variable acoustic heart signals, was able to accurately predict the physiological condition from the signal's morphology. On the first dataset, the instantaneous breathing pressure and the level of breathing resistance were estimated, while in the second dataset, the level of stress, correlated with cardiac contractility, was predicted. With the correct choice of signal representation and analysis parameters, the proposed framework may be applied to continuous non-invasive monitoring of cardiac and respiratory functions, thus providing a new technology for detection and diagnosis of mechanical dysfunctions caused by cardiovascular and cardiopulmonary diseases.

References

[1] R.F. Rushmer, Cardiovascular Dynamics, fourth ed., WB Saunders Co., Philadelphia, 1978.

[2] B. Bromberger-Barnea, Mechanical effects of inspiration on heart functions: a review, Federation Proc. 40 (1981) 2172–2177.

[3] M.E. Tavel, Clinical Phonocardiography & External Pulse Recording, third ed., Year Book Medical Publishers Inc., Chicago, 1978.

[4] T. Sakamoto, et al., Hemodynamic determinants of the amplitude of the first heart sound, Circ. Res. 16 (1965) 45–57.

[5] W.B. Clarke, S.M. Austin, M.S. Pravib, P.M. Griffen, J.T. Dove, J. McCullough, B.F. Schreiner, Spectral energy of the first heart sound in acute myocardial ischemia, Circulation 57 (1978) 593–598.

[6] A.A. Luisada, Sounds and pulses as aids to cardiac diagnosis, Med. Clin. N. Am. 64 (1) (1980) 3–32.

[7] M.E. Tavel, Cardiac auscultation: a glorious past—and it does have a future! Circulation 113 (9) (2006) 1255–1259.

[8] F. Enseleit, F. Duru, Long-term continuous external electrocardiographic recording: a review, Europace 8 (4) (2006) 255–266.

[9] R.E. Kleiger, et al., Decreased heart rate variability and its association with increased mortality after acute myocardial infarction, Am. J. Cardiol. 59 (4) (1987) 256–262.

[10] A.P. Davie, et al., Value of the electrocardiogram in identifying heart failure due to left ventricular systolic dysfunction, BMJ 312 (7025) (1996) 222.

[11] L.G. Durand, P. Pibarot, Digital signal processing of the phonocardiogram: review of the most recent advancements, Crit. Rev. Biomed. Eng. 23 (3–4) (1995) 163–219.

[12] R.M. Rangayyan, R.J. Lehner, Phonocardiogram signal analysis: a review, Crit. Rev. Biomed. Eng. 15 (3) (1988) 211–236.

[13] J.C. Wood, A.J. Buda, D.T. Barry, Time–frequency transforms: a new approach to first heart sound frequency dynamics, IEEE Trans. Biomed. Eng. 39 (7) (1992) 730–740.

[14] M.S. Obaidat, Phonocardiogram signal analysis: techniques and performance comparison, J. Med. Eng. Technol. 17 (6) (1993) 221–227.

[15] D. Chen, et al., Time–frequency analysis of the first heart sound. Part 2: An appropriate time–frequency representation technique, Med. Biol. Eng. Comput. 35 (4) (1997) 311–317.

[16] L.G. Durand, et al., Evaluation of FFT-based and modern parametric methods for the spectral analysis of bioprosthetic valve sounds, IEEE Trans. Biomed. Eng. 33 (6) (1986) 572–578.

[17] G. Amit, N. Gavriely, J. Lessick, N. Intrator, Acoustic indices of cardiac functionality, in: International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS), 2008, 77–83.

[18] L.G. Durand, et al., Comparison of pattern recognition methods for computer-assisted classification of spectra of heart sounds in patients with a porcine bioprosthetic valve implanted in the mitral position, IEEE Trans. Biomed. Eng. 37 (12) (1990) 1121–1129.

[19] Z. Guo, et al., Artificial neural networks in computer-assisted classification of heart sounds in patients with porcine bioprosthetic valves, Med. Biol. Eng. Comput. 32 (3) (1994) 311–316.

[20] P.M. Bentley, P.M. Grant, J.T. McDonnell, Time–frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds, IEEE Trans. Biomed. Eng. 45 (1) (1998) 125–128.

[21] A. Wismüller, O. Lange, D.R. Dersch, G.L. Leinsinger, K. Hahn, B. Pütz, D. Auer, Analysis of biomedical image time-series, Int. J. Comput. Vision 46 (2) (2002) 103–128.

[22] R.G. Stockwell, L. Mansinha, R.P. Lowe, Localization of the complex spectrum: the S-transform, IEEE Trans. Signal Process. 44 (4) (1996) 998–1001.

[23] L. Cohen, Time–frequency distributions—a review, Proc. IEEE 77 (1989) 941–981.

[24] S.C. Johnson, Hierarchical clustering schemes, Psychometrika 2 (1967) 241–254.

[25] J.H. Ward, Hierarchical grouping to optimize an objective function, J. Am. Stat. Assoc. 58 (301) (1963) 236–244.

[26] R.O. Duda, P.E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973.

[27] R. Fisher, The statistical utilization of multiple measurements, Ann. Eugenics 8 (1938) 376–386.

[28] L.G. Durand, et al., Spectral analysis and acoustic transmission of mitral and aortic valve closure sounds in dogs. Part 1. Modelling the heart/thorax acoustic system, Med. Biol. Eng. Comput. 28 (4) (1990) 269–277.

[29] D. Gill, N. Gavriely, N. Intrator, Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model, Comput. Cardiol. (2005).