Morphological Determination of Pathological PCG Signals by Time and Frequency Domain Analysis

Márton Áron Goda[1], Péter Hajas[2]

[1]Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary
[2] Meditech Ltd., Budapest, Hungary

Abstract

With the development of diagnostic devices over the past few decades, the algorithmic classification of heart sound recordings has become possible. Although this field has been under research for a relatively long time, the classification of such recordings is not yet straightforward.

We were given a large manually classified database of heart sounds with the challenge [1]. We worked together with an experienced cardiologist to find the aspects affecting the classifications.

To algorithmically classify a heart sound recording as normal or abnormal, it is necessary in most cases to accurately locate both the fundamental heart sounds and the systolic and diastolic regions. For this purpose we used the method provided in the example entry [2][3]. Minor modifications were made, such as tuning some of the parameters to match the database parameters.

In the classification of the heart sounds, we were looking for the morphological features of the abnormal signals, for example, mitral stenosis, mitral insufficiency, aortic stenosis, aortic insufficiency, tricuspid stenosis and tricuspid insufficiency. We extracted several features from both time and frequency domains, for example, the frequency properties of systolic and diastolic segments and resampled wavelet envelope features.

The extracted features were classified by the help of a support vector machine. In order to train the classifier, we used a reduced, sorted dataset with a more balanced ratio of abnormal and normal signals. During the official phase, our best scores on a random subset were 77.2% sensitivity, 85.2% specificity and 81.2% final modified accuracy (MAcc). Our scores for the entire test dataset are 83.77% sensitivity, 76.8% specificity and 80.28% MAcc. Our scores for the entire training dataset are 93.08% sensitivity, 84.70% specificity and 88.70% MAcc.

Keywords: heart sound analysis, support vector machine, wavelet decomposition, PhysioNet/CinC Challenge 2016

1. Introduction

Nowadays, because of medical and technical advancements, the number of patients in the population who have survived heart attacks is growing continuously [1]. Since more and more patients survive the acute event, the number of patients with heart disease continues to increase.

The phonocardiogram (PCG) is a very simple and cost-effective diagnostic tool for heart sound examination: the cardiologist auscultates the acoustically propagated heart sound signals. In spite of offering only acoustic signal, it is also capable of describing heart disease. Algorithmic processing of phonocardiograms could help to expand the limits of this monitoring tool. As opposed to ECG, which only describes the physiological operation of the heart, PCG can distinguish between different pathological cases.

In applying processing algorithms to PCG sound data, one has to take into account that several factors can influence the characteristics of the recordings [4]. These can be of patient origin (for example stomach or lungs sounds), or of external origin (for example a conversation). The signal recording and processing chain, the type of the sensor, the way it is used and pressed on the skin and the filtering applied (linear and nonlinear) can also considerably influence the signal quality and signal properties [4].

As it is described in detail in [1], individual databases were recorded with different recorders and collected independently. One example is the significantly different frequency components of database *b* and database *a*, according to our experiences. Consequently, by knowing the exact device type one can improve a given algorithm.

Physiological cases have to be distinguished from pathological cases. For example a higher standard deviation in the rhythm, such as arrhythmia, does not necessarily mean any pathological heart failure in the valves and/or cardiovascular system. Of course, the pathological and physiological symptoms mutually interact with each other. On the other hand, these two cases of the heart dysfunction can not be separated, because the heart continuously adapts itself to failure during its whole lifetime.

2.   Preprocessing

The sampling frequency of the recordings in the database is 2 kHz. For processing purposes we were using the downsampled signal with a sampling frequency of 1 kHz. The lower cut-off frequency of most sensors in the given databases was 20 Hz. In order to reduce the effect of the different signal-collecting methods we decided to filter all recordings with a third-order bandpass filter between 20–400 Hz before processing.

As mentioned above, artefacts and noises can corrupt the sound recording, and so can also affect processing. One category of artefacts could be observed in training database *e*: completely silent segments in the PCG signal. If we detected silent segments longer than 25 ms, the longest non-silent region was chosen for further processing.

The next step in the processing chain was to derive cardiac parameters, such as heart rate and systolic length. For this purpose the algorithm provided in the sample entry was used with minor modifications [3]. We calculated the cardiac cycle lengths and the systolic lengths for the training database by using the hand-corrected annotation of the recordings. Heart rate in the training database varied between 35 bpm and 159 bpm, and the systolic interval between 0.172 s and 0.537 s.

As a result, we modified the boundaries to search for the peak, which corresponded to the heart rate in the autocorrelation function. The new boundaries used were 0.4 s and 1.8 s. For systolic peak searching we restricted the upper limit to a maximum of 550 ms.

As a result the number of errors decreased from 112 to 81 in the case of heart rate and from 92 to 47 in the case of systolic length. We regarded an estimation correct if it did not differ from the reference by more than 10% in the case of heart rate and by more than 50 ms in the case of systolic interval.

The next step was the identification of the heart sound boundaries, for which the original Springer algorithm was used [2]. As a result the onset of S1, systole, S3 and diastole were obtained.

3.   Feature extraction

The next step of the signal processing was the extraction of the features. They may fall into three categories:
1. time-domain features,
2. frequency-domain features and
3. wavelet envelope features.

3.1.   Time-domain features

The width of S1 in a recording was calculated as the average of S1 widths in the recording after the upper and lower 10% the outliers had been discarded. The same

process was applied to S2. These features were also used in the sample entry [1].

3.2.   Frequency-domain features

Fourier transform is a widely used tool in the analysis of signals. Using Discrete Fourier Transform (DFT) the discrete time series of the sampled sound signal can be transformed from the time-domain to the frequency-domain. One big drawback is that, since real world signals are not stationary, the time localization of the signal frequencies is lost. We used the DFT algorithm to obtain the mean amplitude spectrum of two of the signal regions: the systolic and diastolic regions. The process was the same for the two regions: it is described below for the systolic region.
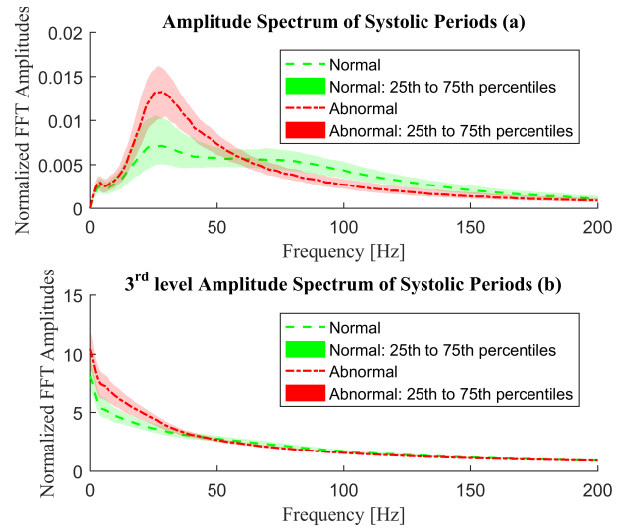

Figure 1.   Mean amplitude spectral statistics for the training database. The central line represents the median of the frequency bins, while the darkened area refers to a range of 25th to 75th percentiles of the bins.

Firstly, the DFT was calculated for each systolic segment in the recording. Filtering and normalization were applied to reduce the effect of noise. The next step was averaging the corresponding frequency bins for the whole recording. After outliers had been rejected, the mean of the amplitude spectra of the segments was calculated, which resulted in a mean amplitude spectrum for the systolic segments of a given record. Statistics from the resulting spectra for the complete training dataset are shown in *Figure 1(a)*.

As features, we used the area under the amplitude spectrum in the frequency ranges, where the two plots were well separable. These frequencies were chosen arbitrarily as follows: 25–35 Hz and 80–120 Hz for the diastolic segment, and 28–35 Hz and 80–120 Hz for the systolic segment.

Some additional features were derived from the amplitude spectrum described above. The normalized mean

amplitude spectrum was transformed twice more using the DFT algorithm, which resulted in the spectra shown in *Figure 1(b)*. As features, we used the area under the amplitude spectrum in the following frequency ranges: 5–25 Hz for the diastolic segment and 8–25 Hz for the systolic segment.

3.3. Wavelet-envelope features

The wavelet transform (WT) is an excellent tool for extracting the main morphological parameters of the phonocardiogram (PCG) [5]. In contrast to Fourier transform, WT can describe both frequency and time domain information.

The characterization of S3 and S4 heart sounds and murmurs is necessary in order to describe the PCG signals morphologically. Since their frequency range does not fully overlap with the frequency ranges of S1 and S2 [1], they can be partly separated using DWT.

In order to extract the third group of features, the following steps were applied to the heart sound recordings:

1. wavelet decomposition,
2. envelope extraction for all levels,
3. envelope resampling for feature vector,
4. dimension reduction (SVD).

In [5], Daubechies wavelet decomposition was used. In their work, a signal with a sampling frequency of 8 kHz was decomposed to five detail and one approximation level. We followed the same method with some modifications. We used 3 detail components (D1–D3) and one approximation component (A3) of the recordings sampled at 1 kHz. The scheme of the decomposition is depicted in *Figure 2*. We found that these components were sufficient for further processing. In [5], the Db4 wavelet basis was used: in this work, we tested Db1, Db4 and Db5 wavelet bases. Although we found that Db5 was the best on the training data, our best score on the test data was obtained by using the Db1 wavelet basis.
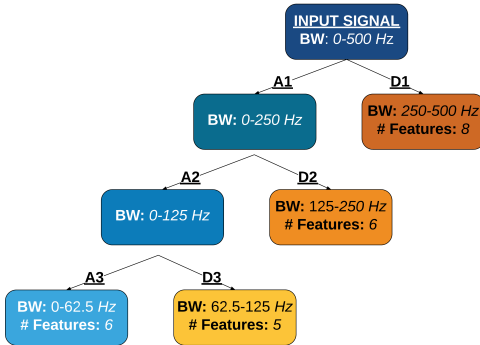

Figure 2. Three-level discrete wavelet decomposition

A method similar to the one described in [4] was used for further processing of the decomposed signal. In [4], a "mean signal focused on the morphological time characteristics of the heart sounds" [4] was extracted. It was calculated as the mean value of the normalized average Shannon energy envelope of the preprocessed components. Then the S1 mean signal was divided into 8 equal parts; for these parts the mean-square value of the envelope was calculated, which resulted in 8 features. Similarly, 24, 8 and 48 features were extracted from systole, S2 and diastole, respectively. We applied a similar method, with two significant differences. Firstly, instead of creating a mean signal and then extracting the features we extracted the features of each segment in each heart cycle and then averaged the features. Secondly, in order to obtain the features in each frequency bands we applied the envelopes on the wavelet decompositions. In addition to the Shannon energy envelope, we tested two more envelopes: the Hilbert amplitude envelope and the Homomorphic envelope. Our best results were obtained by using the Homomorphic envelope.

The extracted feature vector with a length of 192 (*Prior set*) is built up as described in *Table 1*. The feature vector is also shown in *Figures 3(a–d)*, where *Figure 3(a)* corresponds to D1 level.

Level	Decomp. Length	S1	Sys.	S2	Dia.	Prior set	Final set
D1	L/2	12	24	12	48	96	8
D2	L/4	6	12	6	24	48	6
D3	L/8	3	6	3	12	24	5
A3	L/8	3	6	3	12	24	6
D1–A3	L	24	48	24	96	192	25

Table 1. Number of features in the wavelet decompositions of heart sounds, where L refers to the length of the original recording

Singular value decomposition (SVD) was used in order to reduce the dimensionality of the extracted feature vector: 8, 6, 5 and 6–altogether 25–features were derived from D1 to A3 (*Final set* in *Table 1*). The length of the derived feature vector was determined experimentally based on the singular values.

3.4. Heart sound SVM classification

The Support Vector Machine (SVM) is a well-known machine-learning approach that is also available in Matlab®. It has some critical points, such as the input parameters, the kernel function and the construction of the training set. The aspects in the evaluation of our SVM models were the score value described in [1] and the generalization rate which was obtained by performing 10-fold cross validation.

We observed significant differences between training sets: we gained the best scores on training set *e* (even if we excluded similar signals) and very low classification loss. However, training the SVM for training set *b* was the least efficient. The number and ratio of the subjects and records were also different in each data set. Therefore it

	Test set			Training database				
Num	Se(%)	Sp(%)	MAcc(%)	Se(%)	Sp(%)	MAcc(%)	ClassLoss	Features
201	**87.1%**	74.8%	81.0%	**97.1%**	82.8%	**90.0%**	0.1579	32
203	77.2%	85.2%	**81.2%**	91.7%	82.2%	87.0%	0.1548	36
209	74.4%	**87.8%**	81.1%	71.1%	**86.6%**	78.9%	0.1883	60

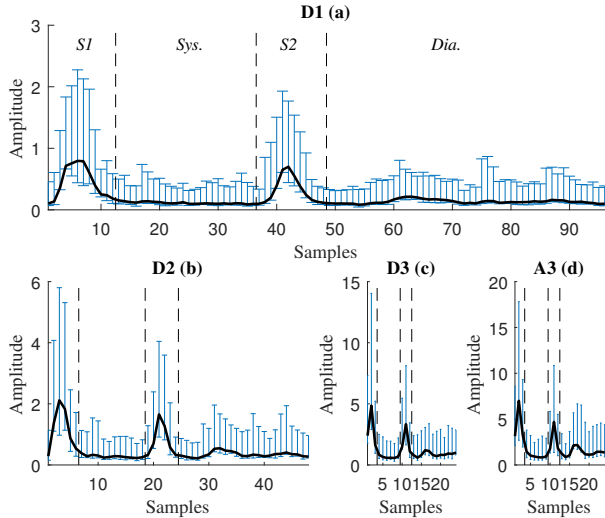Table 2. Results on test and training databases

Figure 3. Resampled wavelet envelopes (black lines). The dashed line separates the heart sound regions (S1, Sys., S2, Dia.). The bars represent the minimum and maximum envelopes in the recording. The morphology of heart sounds in the recording can be observed in different frequency bands.

was not self-evident to compose the ideal training set for the SVM to learn the major range of normal and abnormal recordings.

3.5. Results

In *Table 2* the results of our three entries are shown. In the first two rows the training set contained 1000 randomly chosen recordings: the only restriction was the 1:1 ratio of the abnormal and normal recordings. In the last row the individual databases and diagnoses were taken into consideration. We arbitrarily determined the number of recordings for a specific patient group.

Unfortunately the results, which can be seen in *Table 2* are not easy to analyse, as there were also important differences in the construction of features, such as the number of features, mother wavelets and DFT properties. The maximum running time did not exceed 10% of quota and the average running time was less than 8% of quota in most cases.

4. Conclusions

This paper presents a machine learning approach for classification of heart sound recordings. We have extracted several features in both the time- and frequency-domains.

We have used SVM in order to classify recordings with several training set configurations. Our scores for the entire test dataset were 83.77% sensitivity, 76.8% specificity and 80.28% MAcc. We have achieved 87.0% MAcc for the entire training set with 82.2% specificity and 91.7% sensitivity.

However, with the feature set used in 209 the following results were obtained for training databases *b–f*: 97.3% sensitivity, 86.2% specificity and 91.7% MAcc. In this case SVM was trained for the half of the training databases. Although there are several difficulties in implementing an effective program, these results may enable orientation of non-specialists in diagnosis.

Future research and development could concentrate on the creation of an algorithm that is able to distinguish between the different types of diseases. Noise immunity of the algorithm and its tolerance towards dissimilaties in recording circumstanses should also be improved in the future.

Acknowledgment

References

[1] Liu C, Springer D, et al. An open access database for the evaluation of heart sound algorithms. Physiological Measurement 2016;37(9).

[2] Springer DB, Tarassenko L, Clifford GD. Logistic regression-hsmm-based heart sound segmentation. IEEE Transactions on Biomedical Engineering April 2016; 63(4):822–832. ISSN 0018-9294.

[3] Schmidt S E; Holst-Hansen CGCTESJJ. Segmentation of heart sound recordings by a duration-dependent hidden markov model. Physiological Measurement 04 2010;31.

[4] Maglogiannis I, Loukis E, Zafiropoulos E, Stasis A. Support vectors machine-based identification of heart valve diseases using heart sounds. Computer Methods and Programs in Biomedicine 2009;95.

[5] Uğuz H. Adaptive neuro-fuzzy inference system for diagnosis of the heart valve diseases using wavelet transform with entropy. Neural Computing and Applications 2012; 21(7):1617–1628. ISSN 1433-3058.