Systolic Murmurs Diagnosis Improvement by Feature Fusion and Decision Fusion

Somayeh Akbari  
Image processing and Information Analysis  
Lab.  
Faculty of Electrical and Computer Eng.,  
Tarbiat Modares University, Tehran, Iran  
somayeakbari.s@gmail.com

Hassan Ghassemian  
Image processing and Information Analysis  
Lab.  
Faculty of Electrical and Computer Eng.,  
Tarbiat Modares University, Tehran, Iran  
ghassemi@modares.ac.ir

Zahra Akbari  
Image processing and Information Analysis  
Lab.  
Faculty of Electrical and Computer Eng.,  
Tarbiat Modares University, Tehran, Iran  
zahraakbari@modares.ac.ir

Abstract— Acoustic sound generated by the heart mechanical activity, can provide useful information about the condition of heart valves. The heart sound auscultation is the fundamental tool in the evaluation of the cardiovascular system. The advantage of this method is fast, inexpensive and noninvasive. Due to human auscultatory limitation and non-stationary characteristics of phonocardiogram signals (PCG), diagnosis based on sounds that are heard via a stethoscope is difficult skill, therefore it requires a lot of practice. This study has proposed a biomedical automatic system for classification of PCG signals, which, recorded by a digital stethoscope. In order to extract various characteristics of PCG signals, the power spectrum estimation, wavelet transform (WT) and Mel frequency Cepstrum coefficients (MFCC) have been used in feature extraction step. Features are given to four classifiers: support vector machine (SVM), k-nearest neighbor (k-NN), multilayer perceptron (MLP) and maximum likelihood (ML). The majority voting combination rule is utilized for fusion of different classifiers. The proposed method has been examined on dataset of 90 PCG records containing healthy and three types of cardiac valve diseases (pulmonary stenosis (PS), Atrial Septal Defect (ASD) and Ventricular Septal Defect (VSD)). The experimental results demonstrate that the classifier fusion rule significantly increases the diagnostic accuracy of abnormal PCG. Our proposed method can be used for online classification of PCG in intelligent diagnosis systems.

Keywords— PCG signal; MFCC algorithm; power spectrum estimation; wavelet transform; feature fusion; decision fusion.

I. INTRODUCTION

Cardiovascular disease are known the second leading reason of death in the world since 1985 [1]. Early diagnosis of heart disease has considerable impact in reducing deaths from heart disease. Heart sounds are composed of two main sounds named as the first heart sound (S1) and the second heart sound (S2). The closure of tricuspid and mitral valves creates S1 and S2 occurs during closure of pulmonary and aortic valves [2]. Cardiac disorders are presented in heart sound before the appearance of other symptoms [3,4]. The murmurs may happen in a systolic segment or in a diastolic segment, or in both segments simultaneously depending on the disease. The Phonocardiogram (PCG) technique provides digital recording and storage of heart sound signals with a high intensity and resolution [5]. Therefore, by applying signal processing and pattern recognition techniques on PCG signals, a method could be introduced that would be helpful for physicians to diagnosis murmurs. In recent years, several researches have been performed automatically diagnosis of murmurs. Elamaran et al. introduced a biomedical system to determine heart sound signals using Fast Fourier Transform (FFT) [6]. Time-frequency representative algorithms have also been widely used in heart sounds analysis such as the Short Time Fourier Transform (STFT), the Wigner–Ville Distribution (WVD) and the Wavelet Transform (WT), due to non-stationary nature of PCG signals [7-13]. Jabbari et al. proposed a novel mathematical model of systolic murmurs to extract features. They applied multivariate matching pursuit (MMP) and MLP classifier to distinguish heart systolic murmurs from normal heart sounds [3]. In some of studies, speech processing techniques such as Mel-Frequency Cepstrum Coefficient (MFCC) algorithm have been applied for heart sounds analysis [14-15]. Hamidi et al. introduced the new approach for feature extraction based on fractal dimension [16]. In the present study, we propose an automated heart sound classification system based on fusion of the multiple classifiers. We apply three methods including the power spectrum density (PSD) estimation, the wavelet transform and Mel frequency cepstrum coefficients for extract various features of PCG signals. Then, we utilize SVM, k-NN, MLP and ML classifiers to determine the proper classifier, and the most appropriate classifier is selected for each feature extraction method. For improvement of diagnostic accuracy, we utilize fusion of classifiers outputs using the majority voting rule to identify final label of PCG signals. The block diagram of proposed algorithm is shown in Fig. 1.

The details of this study are organized into six sections. Data collection is introduced in the next section. The processing procedure is explained in section III. Experimental results and discussion are presented in section V. finally, the conclusion is given in last section.
II. DATA COLLECTION & PREPROCESSING

We have acquisitioned PCG and electrocardiogram (ECG) signals of 90 children aged 1 month to 8 years. The dataset included 40 volunteer normal subjects and 53 subjects with systolic murmurs consisting of 12 VSD, 21 ASD and 20 PS samples. PCG and ECG signals have been recorded from the Apex in the supine location for 10 sec. Labels of PCG signals have been determined by a pediatric cardiologist using echocardiography test. The ECG and electronic stethoscope (Welch Allyn Meditron) have been used for parallel ECG and PCG signals recording with sampling frequency of 44.1 kHz and 16 bits’ resolution.

The preprocessing step includes filtering, resampling, normalization and segmentation stages:

Filtering: The frequency components of heart sound signals are in below 1 kHz. So, the Butterworth band pass filter with cutoff frequencies of 15 and 1000 Hz is applied on PCGs and with frequencies of 0.5 and 100 Hz is performed on ECGs for removing high frequency noise and eliminating DC values and motion artifacts.

Resampling: Most energy of heart sound signals is below 1 kHz and also for rising processing speed, heart sounds are down sampled to 2 kHz.

Normalization: To reduce amplitude variations of PCG signals, all PCGs are normalized using (1):

\[
x_{\text{norm}} = \frac{x[n]}{\text{norm}(x[n])}
\]  

where \(n\) is the number of data points, \(x[n]\) is the PCG signal, and \(x_{\text{norm}}\) is the normalized signal.

Segmentation: Segmentation stage involves extracting of heart sound cycles from a continuous PCG signal. A single heart sound cycle comprises four components S1, systole, S2 and diastole [Fig. 2]. In this study, auxiliary signal ECG has been utilized for PCGs segmentation. Fig. 2 shows the correlation between events of ECG and PCG signals. According to the correlation between ECG and PCG, S1 start time is 25 ms before the R peak and the end of S1 is selected in accordance with the desired cycle length and about 140 ms after starting that. Start point of S2 is found based on the maximum place between the start time of T wave and 100 to 150 ms after that in the PCG signal. So, at first, location of R peaks is recognized by Tompkins algorithm [17] and then the boundaries of S1 and S2 are estimated according to some clinical knowledge [3]. By using segmentation algorithm, 1355 heart sound cycles are obtained (including 552 normal cycles, 169 ASD cycles, 276 PS cycles and 357 VSD cycles).

III. PROCESSING PROCEDURES

A. Feature Extraction

To extract proper features of biomedical signals which are obtained in time domain, transferring them to various domains can be useful. In this study, we use spectral and time-frequency features.
1. Spectral analysis

The spectral analysis of heart sound signals can be used as an indicator to detect the difference between several heart murmurs. So, the nonparametric power spectral density estimation algorithm using Fourier transform is utilized to calculate of heart sound spectral characteristics. The PSD is defined as follows:

$$PSD_x(f) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \exp(-j2\pi nf / N)$$ \hspace{1cm} (2)

where, $x(n)$ denotes PCG signal and $N$ is the length of signal. Then, we normalized PSD features, that is given:

$$PSD_{X_{norm}}(f) = \frac{PSD_X(f)}{\text{std}(PSD_X(f))}$$ \hspace{1cm} (3)

where, $\text{std}(PSD_X(f))$ is standard deviation of $PSD_X(f)$. Fig. 3 shows the PSD of PCG signals that are estimated by this approach. PSD frequency components of normal PCG are concentrated on a narrow bound, comparatively PSD of VSD murmur has the high density frequency components which are spread along a wide area. PSD of ASD murmur and PS murmur overlap with PSD of normal PCG but they are different in PSD morphology. After applying normalized PSD, Due to heart sounds frequency range, frequency band of 0-1000 Hz with 4 Hz spectral resolution are selected for features extraction. Thus, 250 features are extracted using this method. Then PCA technique is applied to reduce dimension of features. To Choosing the most appropriate principal components, different numbers of principal components are tested to investigate their effects on classifiers performance.

2. Mel Frequency Analysis

Human auditory perception of sound frequency content does not pursuit a linear scale and follows a Mel frequency scale. Sensitivity of the human auditory system to low frequencies (under 1kHz) is more than higher frequencies (above 1kHz). The MFCC operates similar the human auditory processing system [15]. A related pitch of an actual frequency, $f$, which is measured in Hz, is computed on a Mel scale by:

$$\text{mel}(f) = 2595 \cdot \log(1 + \frac{f}{700})$$ \hspace{1cm} (4)

Calculation steps of MFCC is described following [18]:

Frame blocking: Heart sound signals have nonstationary nature. However, they can be assumed as stationary in a short time frame. Therefore, frame size and shift size are determined from experimental results and frame size is intended 256 samples for each frame with 128 samples overlap between each frame.

Windowing: The Hamming window is applied on each obtained frame to minimize the spectral distortion.

FFT: FFT is used to convert the time domain signal into a frequency domain.

Triangular window filterbank: The triangular filterbank has uniformly distributed at the Mel-warped spectrum. The output energy of each filter is calculated by:

$$m(l) = \sum_{k=\alpha(l)}^{h(l)} W_j(k) |X(k)|^2$$ \hspace{1cm} (5)

where, $W_j(k)$ indicates triangle filter, $X(k)$ is the output of FFT stage and $l=1,2,...,L$ refers the filter order which is chosen $L=24$, $\alpha(l)$ and $h(l)$ specify the lower and upper frequency value of the $l$th triangle filter respectively and $c(l)$ determines central frequency of the $l$th triangle filter. That is computed for 24 filter banks between 0 and 1000 Hz as follows [21]:

$$\Delta'_{\text{mel}} = 700(10^{\frac{2595}{21}} - 1) = 41.66$$ \hspace{1cm} (6)

Discrete Cosine Transform: Finally, DCT is applied on the logarithmic energy of filter output to obtain $i$th MFCC by:

$$C_{\text{mfcc}}(i) = \sum_{l=1}^{L} \log(m(l)) \cos \left( \left(1 - \frac{1}{2} \right) \frac{i\pi}{L} \right)$$ \hspace{1cm} (7)

where, $M$ is the desired number of the cepstrum coefficient which is considered $M=12$. Totally 84 features are extracted for each PCG signal by MFCC algorithm, and PCA are utilized to dimension reduction of features.

3. Wavelet Analysis

Wavelet transform can provide the time-frequency domain information with a multi resolution analysis [19]. Discrete Wavelet transform (DWT) of signal $f(x)$ is defined by:

$$W_{\psi} f(j,k) = 2^{-j/2} \int f(x) \psi(2^{-j} x - k) dx$$ \hspace{1cm} (8)

where $j$ is the scale parameter, inversely proportional to the frequency and $k$ is the shift parameter. The DWT decomposes the signal in two components. Approximation coefficients are obtained using a low-pass filter and detail coefficients are estimated by a high-pass filter. The DWT obtains high frequency resolution at high frequency PCG signals and low resolution frequency for low frequency components. For feature extraction, we experience several mother wavelets with 13 level of decomposition and compute logarithmic energy of DWT coefficients as features using:

$$C = \log \left( \sum_{n=0}^{N} c_n^2 \right)$$ \hspace{1cm} (9)

where, $c_n$ is relative with 1-13th detail coefficients and 13th approximation coefficient.

B. Multiple Classifier System

In recent years, using multiple classifiers systems for diagnosing murmurs have been increased [20-22]. Multiple classifier fusion can improve classification accuracy compared with each of the constituent classifiers accuracy. Among all the fusion algorithm, majority voting rule can be...
implemented simply, and it has provided possible of theoretical analysis [23]. This approach is a class-conscious or static combination method [24].

In this research, k-NN, SVM, MLP and ML classifiers are employed for determine best classifier to each feature extraction methods. According to the obtained results, k-NN classifier is chosen to PSD feature classification and SVM classifier is selected to MFCC an WT features classification. We apply majority voting fusion rule on output of classifiers to improving the diagnosis accuracy of systolic murmurs. The outputs of different individual classifies are combined by the uniform majority voting to give the final prediction. In the other hand, the output of each classifier is 0 or 1, so the final label is typically computed by the maximum membership rule [24].

IV. EXPERIMENTS AND DISCUSSIONS

The proposed algorithm is tested on 93 heart sound records to evaluate its effectiveness for improving the hear sounds diagnosis accuracy. As mentioned in preprocessing section, 1355 PCG cycles are obtained containing 552 normal PCG cycles, 169 ASD cycles, 276 PS cycles and 357 VSD cycles. The 50% of PCG cycles are considered as the training set, and the other 50% are utilized for the testing set. The training and testing sets of all classifiers are the same. The accuracy and the validity have been used for performance evaluation of classifiers.

In classification stage, PSD features are classified using k-NN classifier with cosine distance metric and \( k=4 \) that has been selected experimentally. In Table I, the effect of different numbers of \( p \) principal components of PSD and MFCC is assessed. SVM classifier is basically introduced for binary classification problems. SVM can be utilized on multi-class classification using hierarchical structure. To aim of this purpose, there are two approaches containing one-against-all (OAA) and one-against-one (OOA) algorithms [25]. In this study, we employ multiclass SVMs based on one-against-all algorithm for MFCC and WT features classification. In order to the classification of MFCC features, multi class SVMs with parameter values; \( C=5 \) and \( \gamma=0.03 \) with RBF kernel have been selected. Also, for WT features, multi class SVMs with kernel function RBF has been used with \( C=100 \) and \( \gamma=0.05 \).

<table>
<thead>
<tr>
<th># of Features</th>
<th>PSD</th>
<th>MFCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc (%)</td>
<td>Val(%)</td>
<td>Acc(%)</td>
</tr>
<tr>
<td>3</td>
<td>62.4</td>
<td>61.2</td>
</tr>
<tr>
<td>5</td>
<td>77.5</td>
<td>77.2</td>
</tr>
<tr>
<td>10</td>
<td>87</td>
<td>87.2</td>
</tr>
<tr>
<td>15</td>
<td>88.6</td>
<td>88.6</td>
</tr>
<tr>
<td>20</td>
<td>88.7</td>
<td>88.8</td>
</tr>
<tr>
<td>25</td>
<td>\textbf{89}</td>
<td>\textbf{89.1}</td>
</tr>
<tr>
<td>30</td>
<td>88.4</td>
<td>88.5</td>
</tr>
<tr>
<td>35</td>
<td>87.6</td>
<td>87.7</td>
</tr>
<tr>
<td>40</td>
<td>87</td>
<td>87.2</td>
</tr>
</tbody>
</table>

The effectiveness of different wavelet mothers is assessed on classifier results that is shown in Table II.

Table III shows the accuracy and validity of each classifier and the majority voting fusion respectively. Table I illustrates the effect of \( p \) parameter in PCA algorithm on classification accuracy and validity. So 10 values between 3 and 40, have been examined on data set as amount of \( p \). The best classification results are achieved when 20-40 principal components are considered. Therefore, results of 30 principal components of PSD and 35 principal components of MFCC features are selected for classifiers fusion. Table II determines best wavelet mother. So, coif5 is chosen as wavelet mother and result of that is combined in fusion step.

In this research, we apply three feature extraction algorithms which can extract various characteristics of heart sounds. PSD, MFCC and wavelet logarithmic energy features are used to extraction of spectral features, features based on heart sound acoustic properties and time-frequency features respectively. fusion results of feature extraction algorithms, cover more comprehensive characteristics of heart sound properties. A total accuracy of 92.5% and a total validity of 92.4% are attained using classifiers fusion. The results demonstrate total accuracy of classifiers fusion rule is increased by 4.1% compared with the best performance of individual classifier.

V. CONCLUSION

The proposed scheme is aimed for developing a computer-aided heart murmurs diagnosis system based on automatic segmentation, feature extraction, classification and fusion of the classifiers output. PSD, MFCC and wavelet logarithmic energy features are employed to spectral features, features based on heart sound acoustic properties and time-frequency features extraction respectively. By applying majority voting fusion rule on different classifiers results (k-NN and SVMs) with various features, the diagnostic accuracy increase considerably. Our approach is performed on 30 normal and 53 pathological heart sound signals. Overall, a total accuracy of 92.5% and a total validity of 92.4% are achieved.

In future works, the classification accuracy especially accuracy of ASD diagnosis, can be improved by extracting divers features. Also, proposed approach can be applied on a broader range of symptoms and pathologies.
Fig. 3. Typical power spectral density of (Top) Normal heart sound (Second) ASD murmur (Third) PS murmur (Bottom) VSD murmur

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>ASD</th>
<th>PS</th>
<th>VSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc. (%)</td>
<td>Val. (%)</td>
<td>Acc. (%)</td>
<td>Val. (%)</td>
<td>Acc. (%)</td>
</tr>
<tr>
<td>PSD+KNN</td>
<td>87.7</td>
<td>92.72</td>
<td>82.1</td>
<td>85.2</td>
</tr>
<tr>
<td>MFCC+SVM</td>
<td>94.6</td>
<td>84.9</td>
<td>69.1</td>
<td>83</td>
</tr>
<tr>
<td>WT+SVM</td>
<td>88</td>
<td>81.82</td>
<td>60.4</td>
<td>64.4</td>
</tr>
<tr>
<td>Majority Voting</td>
<td>95.7</td>
<td>90.3</td>
<td>79</td>
<td>89.2</td>
</tr>
</tbody>
</table>
REFERENCES