



# Segmentation-based heart sound feature extraction combined with classifier models for a VSD diagnosis system



Shuping Sun<sup>a,\*</sup>, Haibin Wang<sup>b</sup>, Zhongwei Jiang<sup>a</sup>, Yu Fang<sup>b</sup>, Ting Tao<sup>a</sup>

<sup>a</sup> Department of Mechanical Engineering, Yamaguchi University, Tokiwadai 2-16-1, Ube, Yamaguchi, Japan

<sup>b</sup> School of Electrical and Information Engineering, Xihua University, Chengdu 610039, China

## ARTICLE INFO

### Keywords:

VSD  
Heart sound segmentation  
Classifier models  
SVM  
Envelopes  $E_T$  and  $E_F$   
MWHT

## ABSTRACT

In this paper, boundary curve models for the diagnostic features  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$  are proposed to diagnose ventricular septal defects (VSD), which are generally divided into 3 types: small VSD (SVSD), moderate VSD (MVSD) and large VSD (LVSD). The VSD diagnosis is accomplished in three steps. First, in the time domain, the diagnostic features  $[T_{12}, T_{11}]$ , which are the time intervals between two adjacent first heart sounds (S1) as well as the interval between S1 and the second heart sound (S2), are extracted from the envelope  $E_T$  for the heart sound (HS); in the frequency domain, the envelope  $E_F$  for every cardiac cycle sound that the HS is segmented into, based on a moving windowed Hilbert transform (MWHT), is proposed to extract the diagnostic features  $[F_G, F_W]$ , which are the center of gravity and the frequency width of the frequency distribution. Second, to evaluate the detection ability of the proposed diagnostic features, a classification boundary method based on the support vector machines (SVM) technique is proposed to determine the classifiers to diagnose the VSD sounds. Furthermore, to simplify these classifiers and make them parameterizable, according to their shapes, the least squares method is employed to build ellipse models for fitting the classification boundary curves. Finally, the numerical results based on the ellipse models are introduced for diagnosis of the VSD. Moreover, to validate the usefulness of the proposed method for sounds besides VSD and normal sounds, aortic regurgitation (AR), atrial fibrillation (AF), aortic stenosis (AS) and mitral stenosis (MS) sounds are used as examples to be detected. As a result, the classification accuracies (CA) achieved is 98.4% for the detection of clinical VSD sounds from normal sounds and are 95.1%, 94.8% and 95.0%, respectively, for the detection of clinical SVSD, MVSD, and LVSD among VSD sounds.

© 2013 Elsevier Ltd. All rights reserved.

## 1. Introduction

A ventricular septal defect (VSD), a type of congenital heart disease (Wikipedia, 2013; MedlinePlus, 2013), is a heart malformation present at birth. A VSD is a hole in the part of the septum that separates the ventricles. The hole allows oxygen-rich blood to flow from the left ventricle into the right ventricle, instead of flowing into the aorta and out to the body as it should (Medicine, 2003; Merck, 2013; Wikipedia, 2013). It is estimated that approximately one infant in 500 will be born with a VSD (Medicine, 2003; Wikipedia, 2013). Clinically, a VSD is generally classified into three types, according to the size of the hole (Medicine, 2003): SVSD, whose defect diameter is smaller than 5 mm (denoted as  $\Phi \leq 5$  mm), MVSD ( $5 < \Phi \leq 15$  mm) and LVSD ( $\Phi > 15$  mm). The general methods for diagnosing the VSD disease include chest X-ray, electrocardiography, ECG, clinical auscultation, and so on (Medicine, 2003; Merck, 2013; Wikipedia,

2013). Among these methods, heart auscultation by using a stethoscope is the simplest, and this method is routinely used as an early diagnostic study. Two studies mention the detection of VSD by heart auscultation. However, one study (Bhatikar, DeGroff, & Mahajan, 2005) showed that the achieved accuracy of the discrimination between innocent murmurs and VSD murmurs was 90%, and the other study (Higuchi et al., 2006) verified that VSDs were incorrectly classified as aortic stenosis (AS) or aortic regurgitation (AR). Furthermore, to date, there have not been any reports on detecting SVSD, MVSD and LVSD. In recent years, many researchers proposed some automatic methods for distinguishing heart sound signals by using computer techniques and digital signal processing technology. No matter which methods are used, they can be generally summarized in two-steps: (1) how to extract the heart sound (HS) analysis features using various signal processing methods, and (2) how to diagnose heart disease from the HS. As for the HS analysis methods, they can be generally divided into two approaches: time domain analysis and frequency domain analysis. In the time domain, because every cardiac cycle is usually composed of S1 and S2, these

\* Corresponding author. Tel.: +81 80 5753 5406.

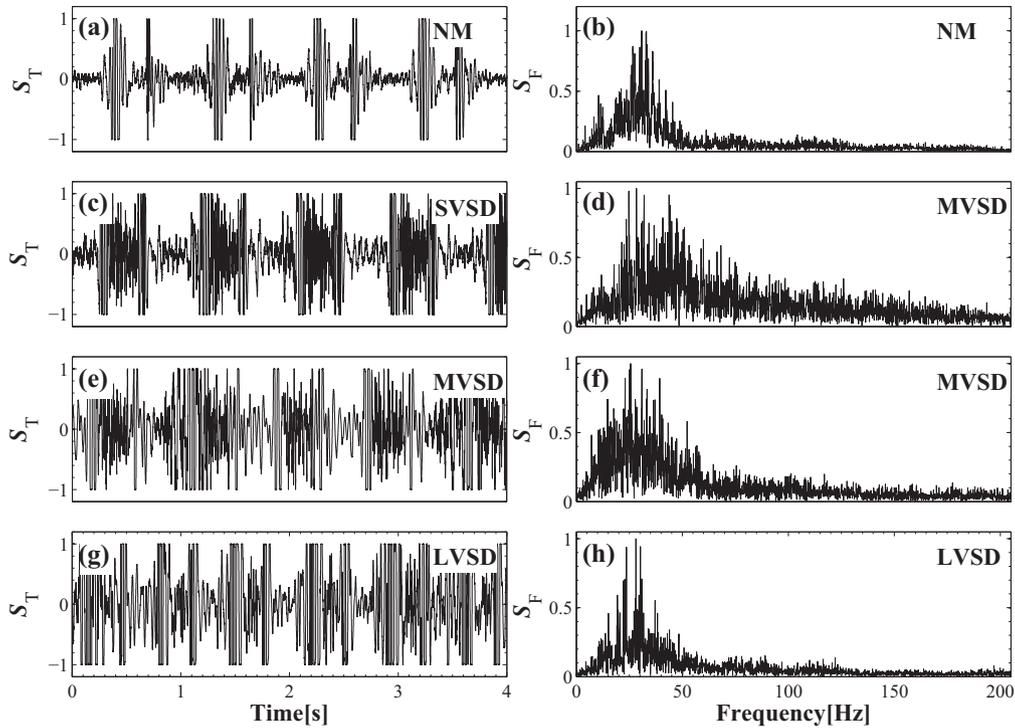
E-mail addresses: [Shp\\_Sun@yeah.net](mailto:Shp_Sun@yeah.net), [sunshp@yeah.net](mailto:sunshp@yeah.net) (S. Sun).

sounds are taken as the emphasis to analyze HS in many studies (Gupta, Palaniappan, Swaminathan, & Krishnan, 2007; Yadollahi & Moussavi, 2006; Martinez-Alajarin & Ruiz-Merino, 2005; Liang, Lukkarinen, & Hartimo, 1997; Syed, Leeds, Curtis, & Guttag, 2006; Jiang & Choi, 2006). Particularly, in one study (Jiang & Choi, 2006), the diagnostic features  $[T_{11}, T_{12}, T_1, T_2]_i$  in the time domain extracted from the characteristic waveform (CSCW) have proven to be useful for the identification of normal sounds and heart diseases, including atrial fibrillation (AF), aortic regurgitation (AR) and mitral stenosis (MS); where  $T_{11}$  is the time interval of two adjacent S1,  $T_{12}$  is the time intervals between S1 and S2,  $T_1$  is the time width of S1, and  $T_2$  is the time width of S2 in  $i$ th sequential HS data. However, unexpected noise is still a difficult problem for this CSCW. Fortunately, another envelope CW (denoted by  $E_T$ ) for HS, based on the Viola integral method, has been proposed and was shown to be effective against not only amplitude variation but complex background noise by the study (Yan, Jiang, Miyamoto, & Wei, 2010). In the frequency domain, many researchers Syed et al. (2006), Wu, Lo, and Wang (1995), Bhatikar et al. (2005), Nygaard, Hasenkam, Pedersen, Paulsen, and Thuesen (1992), Turkoglu, Arslan, and Ilkay (2002), Reed, Fritzon, and Reed (2004), Turkoglu, Arslan, and Ilkay (2003), Travel and Katz (2005), Kim (2003), Sava, McDonnell, and Fox (1994), Iwata, Ishii, Suzumura, and Ikegaya (1980), Kim, Lee, Hub, and Chang (1998) and Adolph, Tanaka, and Stephens (1970) have been concerned with the characteristic extraction by the local frequency analysis method. However, in one previous study (Choi & Jiang, 2010), using two diagnostic features  $[F_{\max}, F_W]$  extracted from the envelope AR-PSD in the frequency domain to detect heart murmurs has been verified as an efficient method because the highest classification accuracies were achieved for classifying normal and abnormal sounds.  $F_{\max}$  describes the maximum peak of the characteristic waveform and  $F_W$  is the frequency width between the crossed points of the characteristic waveform on a selected threshold value. However, it is a pity that there was not a detail explanation about how to select one cardiac cycle HS. Moreover, the method for extracting the envelope in the frequency domain was overly complicated. Recently, a novel method for the selection of the cardiac cycle sound, MWHT, proposed by our studies (Sun et al., 2013) has been reported to provide sufficient performance. The moving average method is a simple way to generate the envelope. Artificial neural networks (ANN) are a computational tool for pattern classification which have been the subject of researchers' interests in the past few years for the classification of heart sounds. For the classification of cardiac sounds, ANN or NN provided a high classification rate and are frequently used as a classifier (Wu et al., 1995; Bhatikar et al., 2005; Nygaard et al., 1992; Turkoglu et al., 2002; Reed et al., 2004; Turkoglu et al., 2003). Recently, support vector machines (SVM) proposed by Vapnik (Cortes & Vapnik, 1995; Vapnic, 1995) have emerged as a new classification technique; SVM have been used successfully for the solution of many problems including heart murmurs classification (Choi & Jiang, 2010), cancer diagnosis (Akay, 2009), handwritten digital recognition (Mehta & Lingayat, 2008), etc. When using SVM and selecting the radio basis function (RBF) as a kernel function, the problems of how to optimize input features  $C$ , which control the tradeoff between complexity of the machine and the number of no separable points, and how to set the best kernel parameter, the width of RBF, are confronted. In a study (Akay, 2009), a proposed grid search approach was verified as an efficient way for searching the parameters. However, the expression of the classifier is so complicated that the calculation for detecting the features is huge. To overcome this problem, a least squares fitting method is widely used in graphic fitting (Gander, Golub, & Strebel, 1994), based on the shape of graphic. According to the above mentioned research, in this study, in the time domain, the envelope  $E_T$  for HS based on the Viola integral method is extracted first. And then in the frequency domain, the envelope  $E_F$  for every cardiac cycle sound

which HS is segmented into, based on MWHT, is extracted. The diagnostic features  $[T_{12}, T_{11}]$  in the time domain and  $[F_G, F_W]$  in the frequency domain are extracted from the  $E_T$  and  $E_F$ , respectively. To evaluate the performance of the proposed method and to build an easily understandable detection system, based on an SVM technique whose parameters are set automatically, the classification boundary curves for data sets (DSs) consisting of  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$  are generated to determine the detecting HS. Furthermore, to simplify these boundary curves based on their shapes, a least squares method was employed to build ellipse models for them. Finally, the efficient of the proposed method is validated by the classification accuracy of the diagnosis abilities for VSD sounds. Moreover, to validate the usefulness of the proposed method, AF, AR, AS and MS sounds are evaluated in addition to normal and VSD sounds.

## 2. Heart sounds auscultation

Auscultation denotes the act of analyzing sounds in the body that are produced in response to mechanical vibrations generated in the organs. The heart sounds can be collected by an electrical stethoscope. In general, there are 4 positions: aortic area, pulmonary area, tricuspid area and mitral area. As for VSD cases, HS collected from the tricuspid area are reported to supply more important information (Bernard Karnath, 2002). In this study, the HS were all collected from the tricuspid area. For every cardiac sound cycle, there are two primary components, S1 and S2, which are generated at the end of atria contraction and the closure of the aortic valve and pulmonary valve, respectively. The original HS signal is denoted here by  $S_T$  in the time domain and its fast fourier transform (FFT) result is denoted as  $S_F$ . For normal sounds, the frequency distribution generally concentrate on the low frequency region. The typical normal sound is named NM and is plotted in Fig. 1. Fig. 1(a) shows the plot of a normal sound in the time domain, which is a sample from a 24-year-old young, healthy man with a heart rate of 68 bpm. Fig. 1(b) is the frequency domain results, where the peak frequency is approximately 30 ~ 40 Hz. For the VSD sounds in patients with a SVSD, there is a minimal shunting of blood and the pressure in the right ventricle remains normal (VSDCause, 2013; Medicine, 2003; Merck, 2013). Because the right ventricular pressure is normal, there is no damage to the lung arterioles and the sound is similar to normal sounds. A typical SVSD sound signal,  $S_T$ , extracted from a female ( $\Phi = 4$  mm, age 3, weight 13 kg, and heart rate 85 bpm) is plotted in Fig. 1(c) and its FFT results is plotted in Fig. 1(d). In patients with a MVSD, the shunting of blood from the left ventricle into the right ventricle is still restrictive. Therefore, it causes insufficient oxygenation of the blood and possibly left heart failure and heart beat is accelerated to a certain degree (VSDCause, 2013; Medicine, 2003). A typical MVSD sound signal,  $S_T$ , extracted from a patient with a MVSD ( $\Phi = 10$  mm) (a female of age 4 with weight 14 kg and 94 bpm heart rate) is plotted in Fig. 1(e) and its FFT results is plotted in Fig. 1(f). In patients with a LVSD, there is a significant shunting of blood from the left ventricle through the right ventricle to the lungs and back to the left atrium and onto the left ventricle, which causes the left atrium and left ventricle to handle an increased amount of blood, and the workload on the heart increases. The increased workload on the heart also increases the heart rate (VSD-Cause, 2013; Medicine, 2003; Merck, 2013). A typical LVSD sound,  $S_T$ , extracted from a male ( $\Phi = 18$  mm, age 0.6, weight 6 kg, and heart rate 124 bpm) is plotted in Fig. 1(g) and its FFT results is plotted in Fig. 1(h). Fig. 1 shows that although the heart rate is getting faster with an increasing VSD size (Figs. 1(a), (b) and (c)), the frequency distribution is getting narrower; although the frequency distribution of LVSD is close to that of NM, the heart rate of LVSD is much faster than that of NM.



**Fig. 1.** Examples for NM and VSD sounds. (a), (c), (e) and (g) show a typical normal and 3 typical VSD sounds in the time domain, (b), (d), (f) and (h) show their fourier transform results.

Therefore, the time domain analysis combined with the frequency analysis might have a better performance than using only the time domain analysis or only the frequency domain analysis.

### 3. Features $[T_{12}, T_{11}]$ and $[F_G, F_W]$ extraction

#### 3.1. HS preprocessing

An HS signal,  $S_T$ , is tested with 16 bit-depth at a sampling frequency  $F_S = 44.1$  kHz. A previous study (Choi & Jiang, 2010) reported that HS signals are mainly dispersed along the frequency range of 20 – 700 Hz. The signal  $S_T$  is then filtered by the wavelet decomposition (WD) method for cancellation of the unwanted frequency components over 689 Hz and below 21.5 Hz. Daubechies 10 wavelet, a good choice as it addresses biomedical signals very well (Nilsson, Funk, Olsson, von Schéele, & Xiong, 2006), is used as the mother wavelet. In the following analysis, the filtered and normalized signal is denoted by  $X[m], 0 = 1, 2, \dots, M - 1$  and is used to extract the envelope  $E_T$  for HS signal.

#### 3.2. $E_T$ extraction

The envelop CSCW proposed in the literature (Jiang & Choi, 2006; Samjin Choi, 2005) has been reported to provide sufficient performance compared to conventional Shannon envelope and Hilbert envelope algorithms, which were used as the empirical or manual way and the automatic selecting way to estimate HS segmentation. However, unexpected noise is still a difficult problem for this method. To overcome this problem, another envelope  $E_T$  for HS, based on the Viola integral method, has been proposed in the study (Yan et al., 2010). This study showed that the  $E_T$  is effective against not only amplitude variation but also a complex background noise. This idea is described in the following

paragraphs. Consider a data series  $X[m]$  preprocessed by WD for a signal  $S_T[m], m = 0, 1, \dots, M - 1$ , where  $M$  denotes the number of the  $S_T$ . In a  $L_T$  neighbourhood of time  $m$ , called the width  $L_T$  time scale, the envelope  $E_T[m]$  is obtained by

$$E_T[m] = \frac{1}{2L_T + 1} \sum_{k=m-L_T}^{m+L_T} (X[k] - \bar{X}[m])^2, \quad m = L_T, L_T + 1, \dots, M - 1 - L_T, \quad (1)$$

where

$$\bar{X}[m] = \frac{1}{2L_T + 1} \sum_{k=m-L_T}^{m+L_T} X[k]. \quad (2)$$

Because the studies (Kumar et al., 2006; Jiang & Choi, 2006) have shown that the duration of S1 or S2 is greater than 0.1 s, in this paper  $L_T = 0.5 \times 0.1 \times F_S = 2205$  is set. Finally, a normalization is applied by setting the maximum amplitude of  $E_T$  to 1.

#### 3.3. HS segmentation

The novel method MWHT proposed by our study (Sun et al., 2013) has been reported to provide sufficient performance compared to a published algorithm (Yan et al., 2010), which was better than the wavelet decomposition method. This idea is described in the following paragraphs. Consider a  $M$ -point discrete-time series  $E_T[m] (m = 0, 1, \dots, M - 1)$ , and suppose the number  $N$  of the moving window  $W_N[l] (l = -(N - 1)/2, -(N - 1)/2 + 1, \dots, (N - 1)/2)$  is an odd number. The  $\tilde{E}_T[n]$  is computed by

$$\tilde{E}_T[n] = \sum_{m=n-\frac{N-1}{2}}^{n+\frac{N-1}{2}} E_T[m] W_N[m-n] W_E \left[ m - \left( n - \frac{N-1}{2} \right) \right], \quad n = \frac{N-1}{2}, \frac{N+1}{2}, \dots, (M-1) - \frac{N-1}{2} \quad (3)$$

where

$$W_E[i] = \begin{cases} \frac{\cos(\frac{N-1-2i}{2N}\pi) - \cos(\frac{N-1-2i}{2}\pi)}{N \sin(\frac{N-1-2i}{2N}\pi)} & \text{for } i = 0, 1, \dots, \frac{N-3}{2}, \frac{N+1}{2}, \dots, N-1, \\ 0 & \text{for } i = \frac{N-1}{2}. \end{cases} \quad (4)$$

According to the best choice in our study (Sun et al., 2013), the window is selected to be a smooth Gaussian window, and the number  $N$  of the Window is set as 1 s.

Based on the t-axis of the positive-to-negative points (PNP) of the  $\tilde{E}_T$  corresponding to the t-axis of the nadirs in the  $E_T$ , the HS can be segmented into every cardiac cycle. The detailed segmentation procedure, shown in Fig. 2, is summarized as follows:

- (a) First, in the time domain, the envelope  $E_T$ , shown as the black lines plotted in Fig. 2(a) for normal sounds and in Fig. 2(b) for VSD sounds, is extracted from the heart sounds  $X$ , shown as the yellow lines plotted in Figs. 2(a) and (b).
- (b) Second, based on Eq. (3), the  $\tilde{E}_T$ s, the black lines plotted in Fig. 2(c) for normal sounds and in Fig. 2(d) for VSD sounds, are generated from the  $E_T$ s plotted in Figs. 2(a) and (b).
- (c) Finally, the PNP, marked by • in Figs. 2(c) and (d), are determined by

$$\text{PNP} = i, \quad \text{if } \begin{cases} \tilde{E}_T(i) = 0, \\ \tilde{E}_T(i-1) \geq 0, \\ \tilde{E}_T(i+1) \leq 0. \end{cases} \quad (5)$$

Then,  $X_{P_i}$ , representing  $i$ th cardiac cycle (Figs. 2(a) and (b)), is calculated by

$$X_{P_i} = X[\text{PNP}_{i+2}] - X[\text{PNP}_i]. \quad (6)$$

Therefore, the HS signal  $X$  is divided into  $X_{P_i} (i = 1, 2, \dots, I)$ , where  $I$  is the number of cardiac cycles signal included in HS signal  $X$ .

### 3.4. $E_F$ extraction

For the  $i$ th cardiac cycle  $X_{P_i}$ , in the frequency domain the envelope  $E_{F_i}$  is obtained by the moving average method as follows:

$$E_{F_i}[k] = \frac{1}{2L_{F_i} + 1} \sum_{l=k-L_{F_i}}^{k+L_{F_i}} |X_{F_i}[l]|, \quad k = L_{F_i}, \dots, N_i - 1 - L_{F_i}, \quad (7)$$

where

$$X_{F_i}[l] = \sum_{n=1}^{N_i-1} X_{P_i}[n] \exp\left(-j \frac{2\pi}{N_i} nl\right), \quad l = 0, 1, 2, \dots, N_i - 1 \quad (8)$$

$N_i$  is the length of  $X_{P_i}$ ,  $|\cdot|$  is the absolute value sign and  $2L_{F_i}$  is the window width. Because the width of the frequency is generally

greater than 16 Hz, then  $L_{F_i} = 0.5 \times 16 \times (F_s)/(N_i)$  is set for the  $i$ th cardiac cycle heart sound  $X_{P_i}$ . Moreover,  $E_{F_i}$  is also normalized.

### 3.5. $[T_{12}, T_{11}]$ and $[F_G, F_W]$ definition

As mentioned in the study (Jiang & Choi, 2006), the time interval between two abutted S1, the interval between S1 and S2, the width of S1, and the width of S2 are very important parameters for detecting heart disorders. A simple way to calculate these intervals is to measure the peaks of S1 and S2. However, the extracted peaks sometimes are not correct, especially for strong heart murmur sounds. To solve this problem in this study, the centers of gravity of S1 and S2 are considered. The concept for defining the diagnostic parameters in the time domain is described in Fig. 3(a).  $H_T$  is the threshold and a suitable value should be selected. The left and right points on the curve  $E_T$  crossed by the  $H_T$  line are defined as  $L_k(i)$  and  $R_k(i) (k = 1, 2; i = 1, 2, \dots, I)$  in a sequential order. The centers of gravity of S1 and S2 in the  $i$ th cardiac cycle  $X_{P_i}$  are defined by  $G_1(i)$  and  $G_2(i)$  and obtained as follows:

$$G_k(i) = \frac{\sum_{m=L_k(i)}^{R_k(i)} m \times E_T^2(m)}{\sum_{m=L_k(i)}^{R_k(i)} E_T^2(m)}, \quad k = 1, 2. \quad (9)$$

So that the time domain features are given by

$$\begin{cases} T_{12}(i) = G_2(i) - G_1(i), \\ T_{11}(i) = G_1(i+1) - G_1(i). \end{cases} \quad (10)$$

On the other hand, the diagnostic features  $[F_{\max}, F_W]$  in the frequency domain have been verified to be useful for detecting heart murmurs (Choi & Jiang, 2010), where  $F_{\max}$  is the frequency at the maximum value and  $F_W$  is the corresponding width of the envelope  $E_F$  over given threshold value, which is described in Fig. 3(b) and and given by

$$F_W(i) = R_F(i) - L_F(i), \quad (11)$$

where  $L_F(i)$  and  $R_F(i)$  are the left and right points of intersection between  $E_{F_i}$  and the given threshold value  $H_F$ . However, the value of  $F_{\max}$ , especially for VSD sounds, will be significantly influenced due to the heart murmurs. Instead of  $F_{\max}$ , the center of gravity of the  $F_G$  for  $E_{F_i}$  is considered as a frequency index, which is described in Fig. 3(b), and is obtained as follows:

$$F_G(i) = \frac{\sum_{l=0}^{[N_i/2]} l \times E_{F_i}^2[l]}{\sum_{l=0}^{[N_i/2]} E_{F_i}^2(l)}. \quad (12)$$

To extract features in the time domain and the frequency domain, the value of  $H_T$  will be a suitable value between the interval  $[0.2, 0.4]$  and the threshold value  $H_F$  has good performance in the interval  $[0.1, 0.2]$ . By experimental analysis, in this paper, the  $H_F$  is

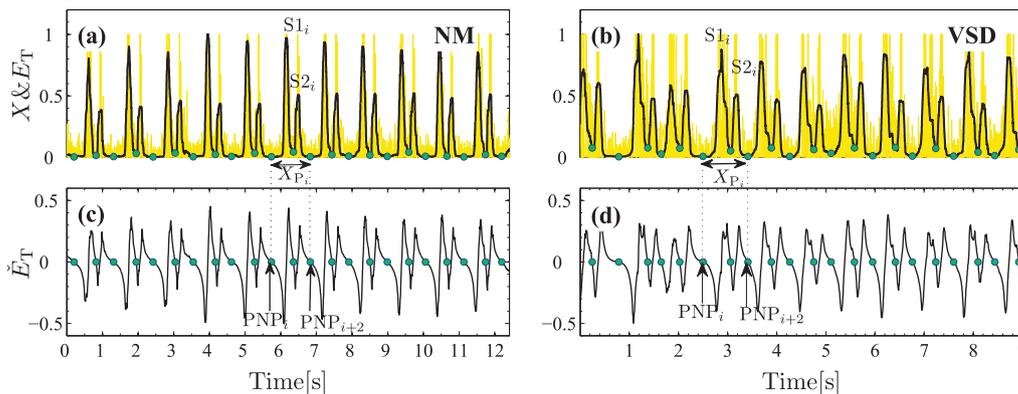
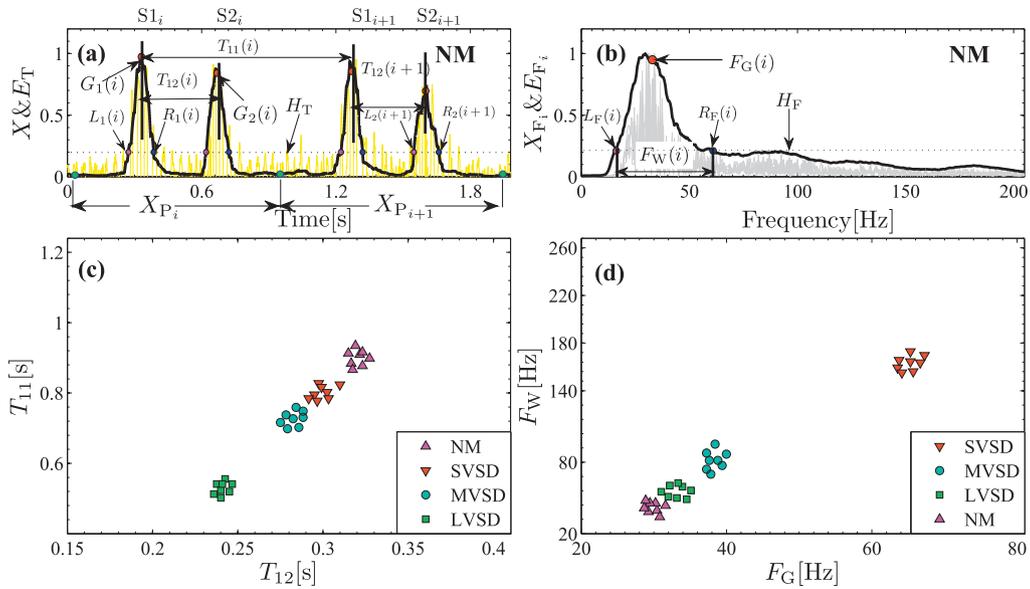


Fig. 2. Example of the segmentation logic for the NM case and VSD cases. The  $\tilde{E}_T$  plotted in (c) and (f) are generated by a Gaussian Window with  $N = 1$  s.



**Fig. 3.** Definition of the diagnostic features  $[T_{12}, T_{11}]$  &  $[F_G, F_W]$  and their scatter diagram. (a) and (b) show a typical normal sound, (c) shows the scatter diagram of  $[T_{12}, T_{11}]$  and (d) shows the scatter diagram of  $[F_G, F_W]$  extracted from the typical NM, SVSD, MVSD, and LVSD sounds.

set at 0.2. Fig. 3(c) and (d) show the plots of  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$ , which are extracted from the typical NM, SVSD, MVSD and LVSD sounds. Fig. 3 shows that

- (1) In the time domain, SVSD is difficult to discriminate from MVSD and NM, but in the frequency domain, SVSD is very easy to discriminate from MVSD and NM.
- (2) In the frequency domain, LVSD is difficult to discriminate from MVSD and NM, but in the time domain, it is easy to discriminate LVSD from MVSD and NM.
- (3) In the frequency domain, the distribution of MVSD is close to NM and LVSD, but in the time domain, MVSD is very far from NM and LVSD.

Therefore, the combination of feature parameters  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$  will be powerful to diagnose VSD.

#### 4. Classification boundary model procedure

In this section, the boundary curves are generated for the DSs consisting of  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$  based on SVM, to simply the classification boundary curve, considering the shape of the classification boundary curve and based on the least squares method; the ellipse models (EMs) of the classification boundary curves, which enclose the DSs with a suitable sensitivity, are built as classifiers to diagnose the VSD sounds.

##### 4.1. Classification boundary function based on SVM

A classification boundary for feature DSs based on the SVM technique is proposed to be used as the classifier to diagnose VSD sounds. To generate a suitable classification boundary for the given feature DSs, other artificially generated DSs, where classification accuracy is based on the automatically searched optimal features, are described as follows.

##### 4.1.1. A review on SVM

SVM have been proposed as an effective statistical learning method for classification of different data classes by the classification curves which are so called support vectors (Vapnik, 1999).

SVM have been used successfully for the solution of many problems, including heart murmur classification (Choi & Jiang, 2010), cancer diagnosis (Akay, 2009), handwritten digital recognition (Mehta & Lingayat, 2008), etc.

The main purpose of SVM is to find a classification boundary curve  $g(x) = 0$ , which has deviated away from all the training DSs and is used to obtain the classification function  $f(x)$ . When the separating samples belong to linear inseparable classes, generally, a non-linear mapping, usually defined as  $\phi(\cdot): R^n \rightarrow R^{nh}$ , is used to map the input vector into a high dimensional feature space. In this case, the classification function  $f(x)$  is defined by

$$f(x) = \text{sign}(g(x)), \quad \text{where } g(x) = W^T \phi(x) + b. \quad (13)$$

Where  $x$  is an input vector,  $W$  is an adjustable weight vector,  $b$  is a bias, and  $g(x)$  is the discriminant function. Suppose the two classes of training DSs consisting of the positive DSs<sup>(+)</sup> and the negative DSs<sup>(-)</sup>. According to the structural risk minimization inductive principle and the Kuhn Tucker optimization theory (Bertsekas, 1995), the approach to solve  $g(x)$  can be written as a classic quadratic optimization problem

$$\max_{\alpha} Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \phi^T(DSs_i) \phi(DSs_j), \quad (14)$$

$$\text{subject to } \begin{cases} \sum_{i=1}^N \alpha_i y_i = 0, \\ 0 \leq \alpha_i \leq C, \end{cases} \quad (15)$$

where  $DSs_i$  is the  $i$ th DSs instance,  $y_i \in \{+1, -1\}$  is a label that determines the class of  $DSs_i$ , and  $C$  is a user-defined positive finite constant. A larger  $C$  means a higher penalty and it usually is assigned to empirical errors. The solution of Eq. (14) should satisfy

$$\alpha_i \left[ y_i \left( \sum_{j=1}^N \alpha_j \phi^T(DSs_j) \phi(DSs_i) + b \right) - 1 \right] = 0, \quad i = 1, 2, \dots, N, \quad (16)$$

which has non-zero multipliers if and only if the points (termed SV) satisfy

$$y_i \left( \sum_{j=1}^N \alpha_j \phi^T(DSs_j) \phi(DSs_i) + b \right) - 1 = 0. \quad (17)$$

The  $g(x)$  is determined by the SV, which is a small subset of the training vectors. Here,  $\phi^T(DSS_i) \cdot \phi(DSS_j)$  can be replaced by a kernel function

$$k(DSS_i, DSS_j) = \phi^T(DSS_i) \cdot \phi(DSS_j). \tag{18}$$

$k(DSS_i, DSS_j)$  may be any of the symmetric functions that satisfy the Mercer conditions (Courant & Hilbert, 1989) and that perform the non-linear mapping into the feature space. In this paper, the Gaussian function is selected as the kernel function.

$$k(DSS_i, DSS_j) = \exp\left(-\frac{\|DSS_i - DSS_j\|^2}{2\tau^2}\right). \tag{19}$$

Hence the classification boundary function of  $x$  can be expressed as

$$g(x, \tau) = \sum_{i=1}^N \alpha_i^* y_i \exp\left(-\frac{\|x - DSS_i\|^2}{2\tau^2}\right) + b^* \tag{20}$$

where

$$\alpha^* = \arg \max \left[ \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j k(DSS_i, DSS_j) \right], \tag{21}$$

$$b^* = y_m - \sum_{i=1}^N \alpha_i^* y_i k(DSS_i, DSS_m), \quad \alpha_m^* \neq 0. \tag{22}$$

Therefore, for a testing vector  $x_s$ , the detection method can be determined by

$$x_s \text{ belongs to } \begin{cases} DSS^{(-)} \text{ class, if } g(x_s, \tau) \leq 0, \\ DSS^{(+)} \text{ class, otherwise.} \end{cases} \tag{23}$$

4.1.2. Classification boundary calculation procedure

To obtain the distribution boundary surrounding a given  $DSS^{(-)}$  with the SVM technique, one needs to build another suitable  $DSS^{(+)}$  which is nearer to or on the distribution boundary  $DSS^{(-)}$ . As an example, suppose  $DSS^{(-)} = [T_{12}, T_{11}]$ , is a DSs whose boundary is to be determined. The data sets  $DSS^{(+)} = [T_{12}^{(+)}, T_{11}^{(+)})$  are then generated by the following algorithm.

- (1) Suppose that  $T_{12}$ , the element of  $DSS^{(-)}$ , follows the normal distribution and calculate its average  $\mu_{T_{12}}$  and standard deviation  $\sigma_{T_{12}}$ .
  - (2) Generate a data set  $T_{12}^{(+)}$  based on the normal distribution with parameters  $\mu_{T_{12}^{(+)}} = \mu_{T_{12}} \pm 8\sigma_{T_{12}}$  and  $\sigma_{T_{12}^{(+)}} = 4\sigma_{T_{12}}$  so that there might be approximately 15% data of  $T_{12}^{(+)}$  overlapped on the boundary of data set  $T_{12}$ .
  - (3) Follow the same process to get  $T_{11}^{(+)}$ . The final data set is obtained as  $DSS^{(+)} = [T_{12}^{(+)}, T_{11}^{(+)})$ .
  - (4) Determine the parameter  $C$  in Eq. (15). Referring to the grid search approach (Akay, 2009), the range of  $C$
- $$C \in 2^{\{-5, -4, \dots, 15\}} \tag{24}$$

is suggested to be an efficient selection. Based on our numerical testing, we set  $C = 2^{15}$ .

- (5) Determine the parameter  $\tau$  in Eq. (19). Based on Chebyshev's inequality,

$$p(|x^{(-)} - \mu_{x^{(-)}}| \leq \tau) \geq 1 - \frac{(\sigma_{x^{(-)}})^2}{\tau^2}. \tag{25}$$

For any probability distribution, if  $\tau$  is set from  $2\sigma_{x^{(-)}}$  to  $4\sigma_{x^{(-)}}$ , there will be at least 75–93.75% of samples close to the mean  $\mu_{x^{(-)}}$ . In our grid search approach program, the kernel parameter  $\tau$  is set as  $\tau \in [2 : 0.1 : 4]\sigma_{x^{(-)}}$ . For each  $\tau$ , one can obtain a boundary curve  $g(x, \tau) = 0$  based on Eq. (20).

- (6) To obtain the optimal parameter  $\tau^{opt}$ , based on Eq. (23) the classification accuracy (CA) for the learning DSs consisting of  $DSS^{(-)} = [T_{12}, T_{11}]^{(-)}$  and  $DSS^{(+)} = [T_{12}, T_{11}]^{(+)}$  at  $\tau \in [2 : 0.1 : 4]\sigma_{x^{(-)}}$  is calculated by

$$CA(\%) = \frac{TP + TN}{TP + FP + FN + TN} \times 100, \tag{26}$$

where TP represents true  $DSS^{(-)} = [T_{12}, T_{11}]$ , TN represents true  $DSS^{(+)} = [T_{12}, T_{11}]^{(+)}$ , FP represents false  $DSS^{(-)} = [T_{12}, T_{11}]$  and FN represents false  $DSS^{(+)} = [T_{12}, T_{11}]^{(+)}$ . The  $\tau$  corresponding to the maximum of CA is selected as  $\tau^{opt}$ .

- (7). The classification boundary function  $g(x, \tau^{opt}) = 0$  is then obtained.

4.1.3. Experimental results for the boundary curves

The total DSs, including 242 normal sound samples (denoted as  $DSS_T^{NM}$  consisting of  $T_{12}$  and  $T_{11}$ , and  $DSS_F^{NM}$  consisting of  $F_G$  and  $F_W$ ) from 16 healthy persons and 226 VSD sound samples ( $DSS_T^{VSD}$  and  $DSS_F^{VSD}$ ), including 62 SVSD sound samples ( $DSS_T^{SVSD}$  and  $DSS_F^{SVSD}$  from 10 patients), 90 MVSD sound samples ( $DSS_T^{MVSD}$  and  $DSS_F^{MVSD}$  from 15 patients) and 74 LVSD sound samples ( $DSS_T^{LVSD}$  and  $DSS_F^{LVSD}$  from 12 patients), are used to obtain the classification boundary functions. Firstly, the 80% DSs are randomly selected from the total DSs to generate the boundary curves  $g_T^{NM} = 0$  and  $g_F^{NM} = 0$  for the normal sounds, and  $g_T^{VSD} = 0$  and  $g_F^{VSD} = 0$  for the VSD sounds. Further, to reduce the influence due to selection of the training data samples, the boundary curves are calculated from three times randomly selected data samples. The results have shown that there is not a big difference between the three obtained boundary curves. The boundary curves shown in Fig. 4 are the averaged curves. Based on Eq. (23), from the features  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$ , the method to detect sound for identifying VSD sounds and normal sounds is defined by:

$$\text{The detecting sound belongs to } \begin{cases} \text{VSD,} & \text{if } g_T^{VSD}([T_{12}, T_{11}]) \leq 0 \text{ and } g_F^{VSD}([F_G, F_W]) \leq 0, \\ \text{NM,} & \text{if } g_T^{NM}([T_{12}, T_{11}]) \leq 0 \text{ and } g_F^{NM}([F_G, F_W]) \leq 0 \\ \text{Not sure,} & \text{otherwise.} \end{cases} \tag{27}$$

To evaluate the performance of these classification boundary curves, by the detection sound Eq. (27) and the elements of the confusion matrix Table 1, classification accuracy (CA), sensitivity (Se) and specificity (Sp) value can be defined as

$$\begin{cases} CA(\%) = \frac{TP+TN}{TP+FP+FN+TN} \times 100, \\ Se(\%) = \frac{TP}{TP+FN} \times 100, \\ Sp(\%) = \frac{TN}{FP+TN} \times 100, \end{cases} \tag{28}$$

where TP represents true VSD sound, TN represents true normal sound, FP represents false VSD sound and FN represents false normal sound. Using the classification boundaries to detect a new DSs consisting of 253 normal and 124 VSD sounds, the classification results summarized in Table 2 show that the sensitivity Se = 98.8%, the specification Sp = 98.1% and the accuracy CA = 98.4%. Because the VSD sounds consist of the SVSD, MVSD and LVSD, 62 SVSD sound samples, 90 MVSD sound samples and 74 LVSD sound samples are used to obtain the boundary curves for SVSD, MVSD and LVSD. The boundary curves are calculated from three times randomly selected data samples. The boundary curves shown in Fig. 5 are the averaged curves. Finally, the new data set from 124 VSD sounds is analyzed and the accuracy Se, Sp and CA are summarized in Table 3.

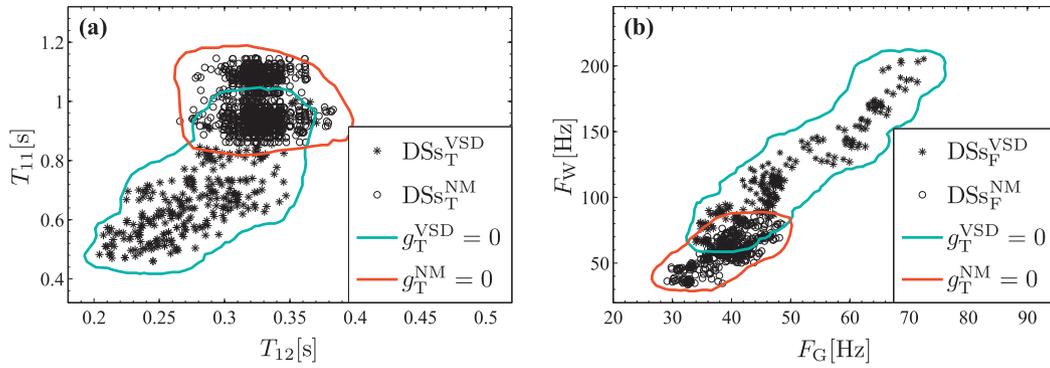


Fig. 4. Experimental results of the classifiers designed for identification of the normal and VSD sounds.

Table 1  
Confusion matrix representation.

Actual	Predicted	
	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Table 2  
Comparative analysis for the classification accuracies by the EMs and boundary curves.

Accuracies (%)	Boundaries	EMs
Se	98.8	98.6
Sp	98.1	98.4
CA	98.4	98.5

4.2. Classification models procedure

4.2.1. Ellipse models

In the above section, although the classification boundary curves (Fig. 4 and Fig. 5) for a training DSs are obtained with higher accuracies by the SVM technique, based on Eq. (20), the mathematical expressions of the classification curves are too complicated to be expressed by parameters. Furthermore, the computation is very intensive. To simplify the classification curves and make them parameterizable, models for the classification boundary curves are considered. Because these shapes of the classification boundary curves plotted in Fig. 4 and Fig. 5 are similar to ellipses, ellipse models for the classification boundary curves are built in this section. Ellipse fitting based on the least squares method is widely

Table 3  
Comparative analysis for the classification accuracies for SVSD, MVSD and LVSD by EMs and Boundary curves.

Accuracies (%)	SVSD		MVSD		LVSD	
	Boundaries	EMs	Boundaries	EMs	Boundaries	EMs
Se	95.2	95.4	94.6	95.1	94.9	95.1
Sp	94.2	95.8	93.1	94.2	93.6	94.8
CA	94.8	95.7	93.7	94.7	94.1	95.0

used (Gander et al., 1994; Chaudhuri, 2010; Prasad, Leung, & Quek, 2012; Ahn, Rauh, & Warnecke, 2001). It is an optimal estimation technique introduced by the maximum likelihood when the random error is assumed to belong to a normal distribution, and it can minimize the error of measurement. Therefore, it can also be seen as a group from the measured value and a group of unknown variables method. An ellipse is a special case of the general conic which can be described by an implicit second order polynomial

$$F(A, x, y) = A_{11}x^2 + A_{21}xy + A_{31}y^2 + A_{41}x + A_{51}y + A_{61} = 0, \quad (29)$$

with an ellipse specific constraint

$$A_{21}^2 - 4A_{11}A_{31} < 0, \quad (30)$$

where  $A_{1i} (i = 1, 2, \dots, 6)$  are coefficients of the ellipse, and  $(x, y)$  is the coordinates of a point lying on it. The polynomial  $F(A, x, y)$  is called the algebraic distance of the point  $(x, y)$  to the given conic. Based on the least squares method, the fitting of a general conic to a set of points  $(x_i, y_i), i = 1, 2, \dots, N$  may be approach by minimizing the sum of the square algebraic distances of the points to the conic which is represented by the coefficient  $A$ :

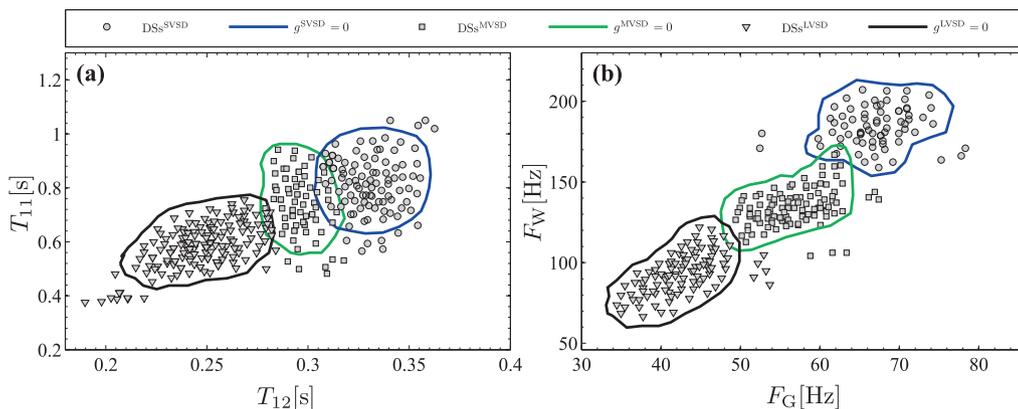


Fig. 5. Classification boundaries and the feature parameters extracted from the SVSD, MVSD and LVSD sounds.

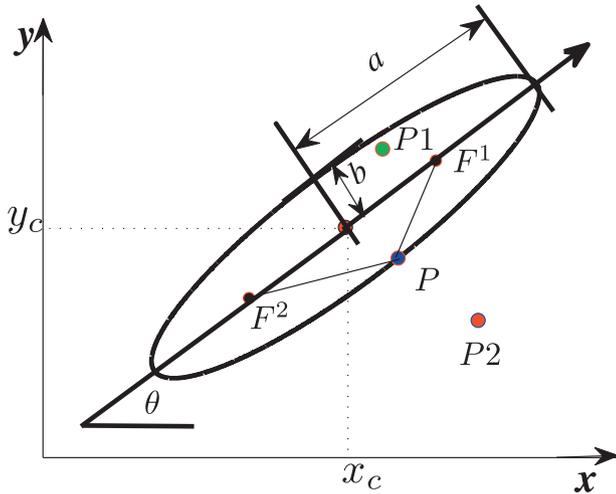


Fig. 6. Ellipse geometric representation in the x-y axis plane.

$$\theta = \begin{cases} 0 & \text{for } A_{21} = 0 \text{ and } A_{11} \leq A_{31}, \\ \frac{\pi}{2} & \text{for } A_{21} = 0 \text{ and } A_{11} > A_{31}, \\ \arccot(\frac{A_{11}-A_{31}}{A_{21}})/2 & \text{for } A_{11} > A_{31} \text{ and } A_{21} \neq 0, \\ \frac{\pi}{2} + \arccot(\frac{A_{11}-A_{31}}{A_{21}})/2 & \text{for } A_{31} > A_{11} \text{ and } A_{21} \neq 0. \end{cases} \quad (36)$$

Therefore, for the points  $P$  on the ellipse,  $P1$  within the ellipse and  $P2$  outside of the ellipse in Fig. 6, according to the definition of the ellipse, the points  $P, P1$  and  $P2$  must be satisfied with

$$\begin{cases} \|P - F^1\|_2 + \|P - F^2\|_2 = 2a, \\ \|P1 - F^1\|_2 + \|P1 - F^2\|_2 < 2a, \\ \|P2 - F^1\|_2 + \|P2 - F^2\|_2 > 2a, \end{cases} \quad (37)$$

where  $\| \cdot \|_2$  represents the Euclidean distance, and  $F^1$  and  $F^2$  are the focus points computed by

$$\begin{cases} F^1 : (x_c + \sqrt{a^2 - b^2} \cos(\theta), y_c + \sqrt{a^2 - b^2} \sin(\theta)), \\ F^2 : (x_c - \sqrt{a^2 - b^2} \cos(\theta), y_c - \sqrt{a^2 - b^2} \sin(\theta)). \end{cases} \quad (38)$$

4.2.2. Ellipse models (EMs) based diagnostic method

The EMs for the classification boundary curves shown in Fig. 4 and Fig. 5 are plotted in Fig. 7 and are expressed as the parameters  $[x_c, y_c, a, b, \theta]$ . The classifier model EMs are  $[0.326, 0.9985, 0.1878, 0.0689, 85.69^\circ]$ ,  $[0.290, 0.7196, 0.0655, 0.3166, 80.3474^\circ]$ ,  $[0.324, 0.825, 0.198, 0.021, 90.6^\circ]$ ,  $[0.296, 0.757, 0.211, 0.020, 90.9^\circ]$ , and  $[0.247, 0.597, 0.171, 0.036, 84.7^\circ]$  for NM, VSD, SVSD, MVSD, and LVSD in the time domain, respectively. In the frequency domain, they are  $[38.7042, 59.2052, 32.046, 8.6774, 76.39^\circ]$ ,  $[54.38, 135.28, 81.68, 9.80, 76.39^\circ]$ ,  $[67.25, 185.49, 28.67, 8.52, 85.7^\circ]$ ,  $[55.08, 132.54, 32.39, 10.07, 76.5^\circ]$  and  $[41.99, 91.82, 33.58, 6.83, 81.29^\circ]$ , respectively. Therefore, by  $[x_c, y_c, a, b, \theta]$  we can determine the distribution of the diagnostic features  $T_{12}, T_{11}, F_C$ , and  $F_W$ . Furthermore, the angle perhaps indicates a pertinence relation between  $T_{12}$  and  $T_{11}$  in the time domain, and between  $F_C$  and  $F_W$  in the frequency domain. Actually, using the ellipse models to diagnose the heart sounds is to judge whether the  $[T_{12}, T_{11}]$  and  $[F_C, F_W]$  are both distributed in the insides of the ellipse models for NM or VSD (SVSD, MVSD, LVSD). According to the relationship between the points and the ellipse using the ellipse definition, the numerical discrimination results denoted as  $NDR_T^{NM}, NDR_T^{VSD}, NDR_T^{SVSD}$  and  $NDR_T^{MVSD}$  in the time domain, and  $NDR_F^{NM}, NDR_F^{VSD}, NDR_F^{SVSD}$  and  $NDR_F^{MVSD}$  in the frequency domain are determined as follows. Suppose the points  $P_T([T_{12}, T_{11}])$  and  $P_F([F_C, F_W])$  are extracted from one heart sound. Here the ellipse models  $EM_T^{MVSD}$  and  $EM_F^{MVSD}$  are taken as the examples to check whether  $P_T([T_{12}, T_{11}])$  and  $P_F([F_C, F_W])$  are inside the ellipse models (Fig. 8). According to

$$\min_A \sum_{i=1}^N F^2(A, x_i, y_i), \quad (31)$$

which can be solved directly by the least squares approach. In this paper, to describe the significance of the ellipse model for the classification boundary curves, the ellipse, as shown in Fig. 6, is presented as the geometric parameters  $[x_c, y_c, a, b, \theta]$ , where the point  $[x_c, y_c]$  is the center,  $a$  is the semi-major length,  $b$  is the semi-minor length, and  $\theta$  is the counterclockwise angle of rotation from the x-axis and the major axis of the ellipse. The  $[x_c, y_c, a, b, \theta]$  corresponding to the ellipse in Eq. (29) is transformed by

$$x_c = \frac{A_{21}A_{51} - 2A_{31}A_{41}}{4A_{11}A_{31} - A_{21}^2}, \quad (32)$$

$$y_c = \frac{A_{21}A_{41} - 2A_{11}A_{51}}{4A_{11}A_{31} - A_{21}^2}, \quad (33)$$

$$a = \sqrt{\frac{2(A_{11}A_{51}^2 + A_{31}A_{41} + A_{21}^2A_{61} - A_{21}A_{41}A_{51} - 4A_{11}A_{31}A_{61})}{(A_{21}^2 - 4A_{11}A_{31})[\sqrt{(A_{11} - A_{31})^2 + A_{21}^2} - A_{11} - A_{31}]}}}, \quad (34)$$

$$b = \sqrt{\frac{2(A_{11}A_{51}^2 + A_{31}A_{41} + A_{21}^2A_{61} - A_{21}A_{41}A_{51} - 4A_{11}A_{31}A_{61})}{(A_{21}^2 - 4A_{11}A_{31})[-\sqrt{(A_{11} - A_{31})^2 + A_{21}^2} - A_{11} - A_{31}]}}}, \quad (35)$$

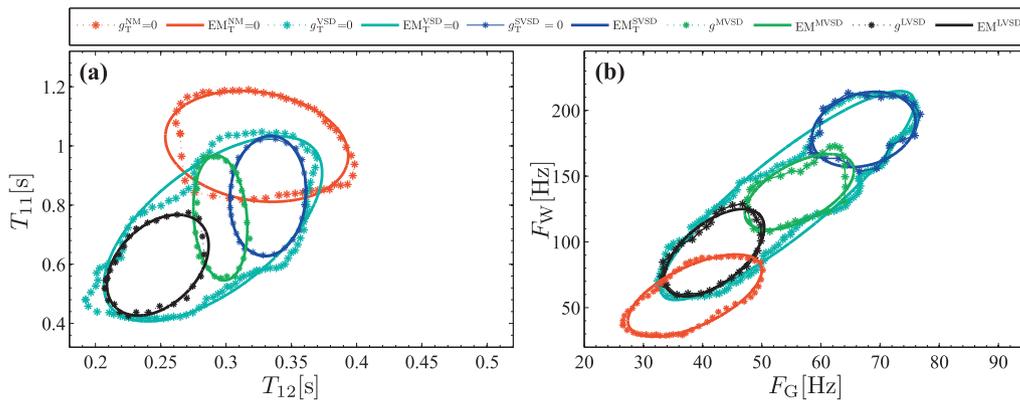
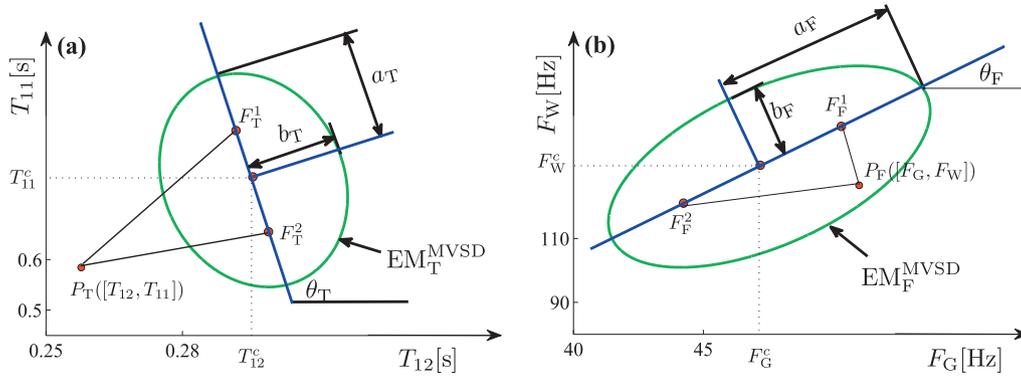


Fig. 7. The classification boundary curves and their ellipse models for normal sounds and VSD sounds. (a) In the time domain and (b) in the frequency domain.



**Fig. 8.** An example for the MVSD ellipse model. (a) In the time domain and (b) in the frequency domain.

the ellipse definition and classification, the  $NDR_T^{MVSD}(P_T)$  can be defined as

$$NDR_T^{MVSD}(P_T) = \begin{cases} 1 & \text{if } \|P_T - F_T^1\|_2 + \|P_T - F_T^2\|_2 \leq 2a_T, \\ -1 & \text{if } \|P_T - F_T^1\|_2 + \|P_T - F_T^2\|_2 > 2a_T, \end{cases} \quad (39)$$

where  $F_T^1$  and  $F_T^2$  are the foci of the ellipse model  $EM_T^{MVSD}$  plotted in Fig. 8(a), and their coordinates are computed by

$$\begin{cases} F_T^1 : (T_{12}^c + \sqrt{a_T^2 - b_T^2} \cos(\theta_T), T_{11}^c + \sqrt{a_T^2 - b_T^2} \sin(\theta_T)), \\ F_T^2 : (T_{12}^c - \sqrt{a_T^2 - b_T^2} \cos(\theta_T), T_{11}^c - \sqrt{a_T^2 - b_T^2} \sin(\theta_T)), \end{cases} \quad (40)$$

where  $(T_{12}^c, T_{11}^c)$  is the center,  $a_T$  is the semi-major axis and  $b_T$  is the semi-minor axis of the ellipse model  $EM_T^{MVSD}$ ; then the  $NDR_F^{MVSD}(P_F)$  can be defined as

$$NDR_F^{MVSD}(P_F) = \begin{cases} 1 & \text{if } \|P_F - F_F^1\|_2 + \|P_F - F_F^2\|_2 \leq 2a_F, \\ -1 & \text{if } \|P_F - F_F^1\|_2 + \|P_F - F_F^2\|_2 > 2a_F, \end{cases} \quad (41)$$

where  $F_F^1$  and  $F_F^2$  are the foci of the ellipse model  $EM_F^{MVSD}$  plotted in Fig. 8(b), and their coordinates are computed by

$$\begin{cases} F_F^1 : (F_G^c + \sqrt{a_F^2 - b_F^2} \cos(\theta_F), F_W^c + \sqrt{a_F^2 - b_F^2} \sin(\theta_F)), \\ F_F^2 : (F_G^c - \sqrt{a_F^2 - b_F^2} \cos(\theta_F), F_W^c - \sqrt{a_F^2 - b_F^2} \sin(\theta_F)), \end{cases} \quad (42)$$

where  $(F_G^c, F_W^c)$  is the center,  $a_F$  is the semi-major axis and  $b_F$  is the semi-minor axis of the ellipse model  $EM_F^{MVSD}$ . In the same way, the  $NDR_T^{NM}(P_T)$ ,  $NDR_T^{SVSD}(P_T)$  and  $NDR_T^{LVSD}(P_T)$  of the point  $P_T$  in the ellipses in the time domain can be determined, and the  $NDR_F^{NM}(P_F)$ ,  $NDR_F^{SVSD}(P_F)$ ,  $NDR_F^{LVSD}(P_F)$ ,  $NDR_T^{LVSD}(P_F)$  of the point  $P_F$  in the ellipses in the frequency domain can be determined. The detection result (DR) for diagnosing the heart sounds by the parameters  $P_T([T_{12}, T_{11}])$  and  $P_F([F_G, F_W])$  is defined by

$$DR = \begin{cases} \mathbf{NM} & \text{for } NDR_T^{NM} + NDR_F^{NM} = 2, \\ \mathbf{VSD} & \text{for } NDR_T^{VSD} + NDR_F^{VSD} = 2, \\ \mathbf{SVSD} & \text{for } NDR_T^{SVSD} + NDR_F^{SVSD} = 2, \\ \mathbf{MVSD} & \text{for } NDR_T^{MVSD} + NDR_F^{MVSD} = 2, \\ \mathbf{LVSD} & \text{for } NDR_T^{LVSD} + NDR_F^{LVSD} = 2, \end{cases} \quad (43)$$

#### 4.2.3. Experimental analysis

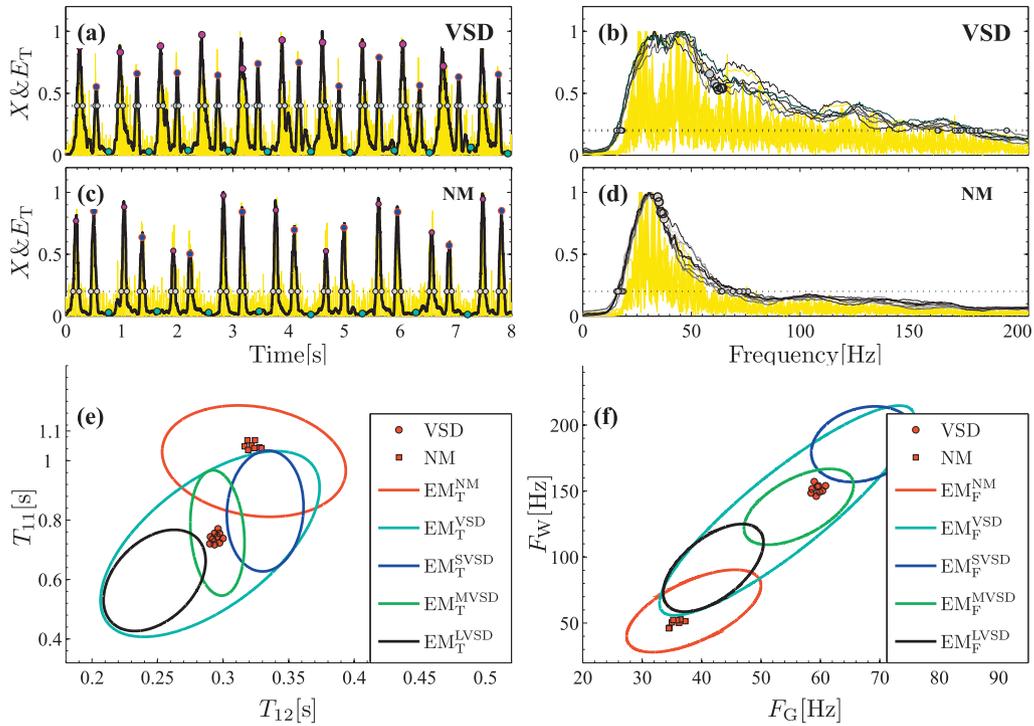
By these EMs and the boundary curves to detect the same DSs used in Section 4.1.3, the comparative accuracy results for NM and VSD sounds are shown in Table 2. The comparative accuracy results for SVSD, MVSD and LVSD are shown in Table 3. The results in.

- (1). Table 2 shows that the accuracy results for NM and VSD are not obviously different between using the ellipse models and the boundary curves.
- (2). Table 3 shows that the performance of the ellipse models seems to be a little better than the boundary curves.

Therefore, the ellipse models can replace the boundary curves to diagnose heart sounds. To explain this detection procedure much better, the VSD sounds (VSD) and normal sounds (NM) are taken as the examples to be analyzed (Fig. 9). Fig. 9(a) shows the VSD sounds which is collected from a female patient ( $\Phi = 12$  mm) of age 1 with 7.5 kg weight. A set of the diagnostic parameters  $[T_{12}, T_{11}]$  extracted at  $H_T = 0.4$  and  $[F_G, F_W]$  extracted at  $H_F = 0.2$  are marked by ● and plotted in Fig. 9(e) and (f). It is obvious that the plots of the parameters  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$  are both concentrated into the ellipse models of VSD and MVSD. Therefore, this sound might be discriminated as VSD and MVSD, which corresponds to the clinical diagnosis because  $\Phi = 12$  mm. Fig. 9(c) shows the normal sound collected from a healthy woman of age 22 with weight 49 kg. A set of the diagnostic parameters  $[T_{12}, T_{11}]$  extracted at  $H_T = 0.2$  and  $[F_G, F_W]$  extracted at  $H_F = 0.2$  are marked by ■ and also plotted in Fig. 9(e) and (f). It is obvious that the plots of the parameters  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$  are both concentrated into the ellipse models of NM. Therefore, the VSD case might be discriminated as normal sound. Therefore, this diagnostic method can easily help the user to understand the detected heart sound in the time domain and in the frequency domain.

#### 4.2.4. Clinical cases analysis

To evaluate the efficiency of the proposed method, besides the NM sounds from three healthy males in Yamaguchi University (in Japan) and two females in Xihua University (in China) and the VSD sounds from patients only with VSD in the Department of Cardiology Surgery of the Chengdu Military General Hospital of PLA (in China), an AR, an AF, an AS, and a MS sounds are randomly selected to be diagnosed. The VSD sounds include two LVSD sounds, which are from a female patient of age 2.2 with weight 6.5 kg (named LVSD1 ( $\Phi = 19$  mm)) and a male of age 2.5 with weight 7.2 kg (LVSD2 ( $\Phi = 17$  mm)); three MVSD sounds, which are from a female of age 0.5 with weight 5.6 kg (MVSD1 ( $\Phi = 14$  mm)), a female of age 2 with weight 11 kg (MVSD2 ( $\Phi = 10$  mm)), and a female of age 3 with weight 13 kg (MVSD3 ( $\Phi = 7$  mm)); two SVSD sounds, which are from a female patient of age 0.4 with weight 8 kg (SVSD1 ( $\Phi = 5$  mm)) and a male of age 1.8 with weight 5.8 kg (SVSD2 ( $\Phi = 5$  mm)). The four normal sounds are collected from a female of age 23 with weight 56 kg (NM1), a male of age 22 with weight 68 kg (NM2), a male of age 27 with weight 75 kg (NM3), and a healthy male of age 18 with weight 70 kg (NM4).



**Fig. 9.** Examples for the heart sound detection procedure. (a) The envelope  $E_T$  and (b) the envelope  $E_F$  of the VSD. (c) The envelope  $E_T$  and (d) the envelope  $E_F$  of the NM. The corresponding diagnostic graphic representations of  $[T_{12}, T_{11}]$  obtained from VSD at  $H_T = 0.4$  and from NM at  $H_T = 0.2$  are displayed in (e) and  $[F_G, F_W]$  obtained from VSD at  $H_F = 0.2$  and from NM at  $H_F = 0.2$  are shown in (f).

The AR, AF, AS, and MS sounds collected from an online clinical training web site are denoted AR, AF, AS, and MS, respectively. The features extracted from LVSD1, LVSD2, MVSD1, MVSD2, MVSD3, SVSD1, SVSD2, NM1, NM2, NM3, NM4, AR, AF, AS, and MS are shown in Fig. 10, and the sound discrimination results are shown as Table 4. Here, MVSD1 is taken as an example to introduce the performance using  $[T_{12}, T_{11}]$ s and  $[F_G, F_W]$ ; the results of MVSD1 in Fig. 10 show that

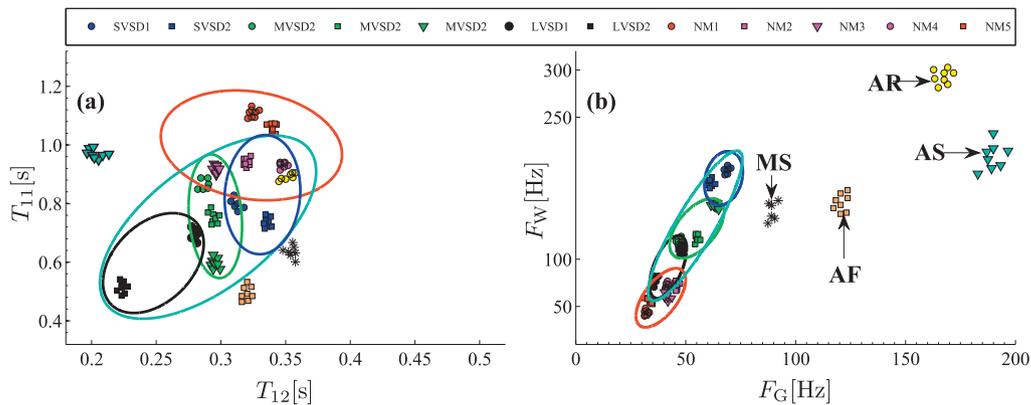
$$MVSD1: \begin{cases} \text{the plots of } [T_{12}, T_{11}] \text{ are distributed within : VSD, MVSD and NM,} \\ \text{the plots of } [F_G, F_W] \text{ are distributed within : VSD, MVSD and LVSD.} \end{cases} \quad (44)$$

Therefore, MVSD1 might be diagnosed as in the NM, VSD and MVSD classes using only  $[T_{12}, T_{11}]$ , and using only  $[F_G, F_W]$  it might be diagnosed as the VSD, LVSD and MVSD classes. However, using  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$  it might be diagnosed as VSD and MVSD, which corresponds to the clinical diagnosis and its detection results denoted as

VSD: MVSD, shown in Table 4. Of course, sound is very difficult to be diagnosed, and problems will occur in this diagnostic system, like for NM2. The results of NM2 in Fig. 10 show that

$$NM2: \begin{cases} \text{the plots of } [T_{12}, T_{11}] \text{ are distributed within : NM, VSD and SVSD,} \\ \text{the plots of } [F_G, F_W] \text{ are distributed within : NM, VSD and LVSD.} \end{cases} \quad (45)$$

Therefore, NM2 might be diagnosed as in the NM and VSD classes (denoted as NM:VSD), which is in keeping with the diagnosis result that the classification accuracy between NM and VSD is 98.4% not 100%. Similarly, the detection results corresponding to Fig. 10 are summarized in Table 4. Furthermore, the analysis results for the typical clinical AR, AS, MR, AF sounds plotted in Fig. 10 show that they are easy to diagnose, and are not mistaken as VSD sounds, which are contrary to the results that the VSD was incorrectly classified as aortic stenosis (AS) or aortic regurgitation (AR), as



**Fig. 10.** Classification models and the distribution of heart sound feature parameters.

**Table 4**

Discrimination results from the numerical discrimination results.

HS	NM		VSD		LVSD		MVSD		SVSD		DR
	NDR <sub>T</sub>	NDR <sub>F</sub>	DNR <sub>T</sub>	NDR <sub>F</sub>							
LVSD1( $\Phi = 19$ mm)	-1	-1	1	1	1	1	-1	1	-1	-1	VSD:LVSD
LVSD2( $\Phi = 17$ mm)	-1	1	1	1	1	1	1	-1	-1	-1	VSD:LVSD
MVSD1( $\Phi = 14$ mm)	1	-1	1	1	-1	1	1	1	-1	-1	VSD:MVSD
MVSD2( $\Phi = 10$ mm)	-1	-1	1	1	-1	-1	1	1	-1	-1	VSD:MVSD
MVSD3( $\Phi = 7$ mm)	-1	-1	1	1	-1	-1	1	1	-1	1	VSD:MVSD
SVSD1( $\Phi = 5$ mm)	-1	-1	1	1	-1	-1	1	-1	1	1	VSD:SVSD
SVSD2( $\Phi = 5$ mm)	-1	-1	1	1	-1	-1	-1	-1	1	1	VSD:SVSD
NM1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	NM
NM2	1	1	1	1	-1	1	-1	-1	1	-1	NM:VSD
NM3	1	1	1	1	-1	-1	-1	-1	1	-1	NM:VSD
NM4	1	1	1	-1	-1	-1	1	-1	-1	-1	NM
AF	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-
AR	1	-1	1	-1	-1	-1	-1	-1	1	-1	-
AS	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-
MS	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-

summarized in Higuchi et al. (2006). Therefore, the proposed method might be efficient to discriminate VSD sounds.

## 5. conclusion

This study proposed the feature extraction and classification boundary curve model diagnostic method, which was useful for detecting VSD and three types of VSD sounds (SVSD, MVSD and LVSD). In the time domain, the features  $[T_{12}, T_{11}]$  were extracted from the envelope  $E_T$  for HS based on the Viola integral method. In the frequency domain, the features  $[F_G, F_W]$  were extracted from the envelope  $E_F$  for every cardiac cycle sound, which was segmented from heart sounds based on a novel MWHT method. To evaluate the detection ability and build an easily understandable detection system for the features  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$ , the boundary curves based on the SVM technique were generated for the DSs consisting of  $[T_{12}, T_{11}]$  and  $[F_G, F_W]$ . Furthermore, to simplify these boundary curves, based on their shapes, a least squares method was employed to build ellipse models for curves. Finally, the efficiency of the proposed method was validated by the classification accuracy of the discrimination abilities for VSD sounds. Moreover, to validate the usefulness of the proposed method, besides normal and VSD sounds, AF, AR, AS and MS heart sounds were evaluated. Therefore, the proposed method perhaps provided an efficient way to obtaining the efficient range of features which characterized different types of heart sounds.

## References

- Ahn, S. J., Rauh, W., & Warnecke, H.-J. (2001). Least-squares orthogonal distances fitting of circle sphere ellipse hyperbola and parabola. *Pattern Recognition*, 34, 2283–2303.
- Akay, M. F. (2009). Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Systems with Applications*, 36, 3240–3247.
- Bernard Karnath, W.T. (2002). Auscultation of the heart. <[http://www.turner-white.com/pdf/hp\\_sep02\\_heart.pdf](http://www.turner-white.com/pdf/hp_sep02_heart.pdf)>.
- Bertsekas, D. (1995). *Nonlinear programming. Optimization and neural computation series*. Athena Scientific.
- Bhatikar, S. R., DeGroff, C., & Mahajan, R. L. (2005). A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics. *Artificial Intelligence in Medicine*, 33, 251–260.
- Chaudhuri, D. (2010). A simple least squares method for fitting of ellipses and circles depends on border points of a two-tone image and their 3-d extensions. *Pattern Recognition Letters*, 31, 818–829.
- Choi, S., & Jiang, Z. (2010). Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique. *Computers in Biology and Medicine*, 40, 8–20.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. In *Machine learning* (pp. 273–297).
- Courant, R., & Hilbert, D. (1989). *Methods of mathematical physics. Number 1*  $\tilde{A}$  in classics library. Wiley.
- Kim, M. D. (2003). Assessment of severity of aortic stenosis through time-frequency analysis of murmur. *Chest*, 124, 1638–1644.
- Gander, W., Golub, G., & Strebel, R. (1994). Least-squares fitting of circles and ellipses. *BIT Numerical Mathematics*, 34, 558–578.
- Gupta, C. N., Palaniappan, R., Swaminathan, S., & Krishnan, S. M. (2007). Neural network classification of homomorphic segmented heart sounds. *Appl. Soft Comput.*, 7, 286–297.
- Higuchi, K., Sato, K., Makuuchi, H., Furuse, A., Takamoto, S., & Takeda, H. (2006). Automated diagnosis of heart disease in patients with heart murmurs: Application of a neural network technique. *Journal of Medical Engineering and Technology*, 30, 61–68.
- Nygaard, H., Thuesen, L., Hasenkam, J. M., Pedersen, E. M., & Paulsen, P. K. (1992). Non-invasive evaluation of aortic valve stenosis by spectral analysis of heart murmurs. In *Proceedings of the Annual International Conference of the IEEE*, 6 (pp. 2584–2585).
- Turkoglu, I., Arslan, A., & Ilkay, E. (2002). An expert system for diagnosis of the heart valve diseases. *Expert Systems with Applications*, 23, 229–236.
- Turkoglu, I., Arslan, A., & Ilkay, E. (2003). An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks. *Computers in Biology and Medicine*, 33, 319–331.
- Iwata, A., Ishii, N., Suzumura, N., & Ikegaya, K. (1980). Algorithm for detecting the first and the second heart sounds by spectral tracking. *Medical and Biological Engineering and Computing*, 18, 19–26.
- Jiang, Z., & Choi, S. (2006). A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope. *Expert Systems with Applications*, 31, 286–298.
- Kim, S. H., Lee, H. J., Hub, J. M., & Chang, B. C. (1998). Spectral analysis of heart valve sound for detection of prosthetic heart valve diseases. *Yonsei Medical Journal*, 39, 302–308.
- Kumar, D., Carvalho, R., Antunes, M., Gil, R., Henriques, J., & Eugenio, L. (2006). A new algorithm for detection of s1 and s2 heart sounds. In *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006 Proceedings (p. II)*, Vol. 2.
- Liang, H., Lukkarinen, S., & Hartimo, I. (1997). Heart sound segmentation algorithm based on heart sound envelopegram. In *Computers in Cardiology* (pp. 105–108).
- Nilsson, M., Funk, P., Olsson, E. M. G., von Schéele, B., & Xiong, N. (2006). Clinical decision-support for diagnosing stress-related disorders by applying psychophysiological medical knowledge to an instance-based learning system. *Artificial Intelligence in Medicine*, 36, 159–176.
- Martinez-Alajarin, J., & Ruiz-Merino, R. (2005). Efficient method for events detection in phonocardiographic signals. In: R.A. Carmona, & G. Linan-Cembrano (Eds.), *Proceedings of SPIE, SPIE 5839* (pp. 398–409).
- Medicine (2013). Ventricular septal defect (vsd). <[http://www.medicinenet.com/ventricular\\_septal\\_defect/article.htm](http://www.medicinenet.com/ventricular_septal_defect/article.htm)>.
- MedlinePlus (2013). Types of congenital heart defects. <<http://www.nlm.nih.gov/health/health-topics/topics/chd/types.html>>.
- Mehta, S., & Lingayat, N. (2008). Svm-based algorithm for recognition of qrs complexes in electrocardiogram. *IRBM*, 29, 310–317.
- Merck (2013). Ventricular septal defect (vsd). <[http://www.merckmanuals.com/professional/pediatrics/congenital\\_cardiovascular\\_anomalies/ventricular\\_septal\\_defect\\_vsd.html?qt=&sc=&alt=>](http://www.merckmanuals.com/professional/pediatrics/congenital_cardiovascular_anomalies/ventricular_septal_defect_vsd.html?qt=&sc=&alt=>)>.
- Travel, M. E., & Katz, H. (2005). Usefulness of a new sound spectral averaging technique to distinguish an innocent systolic murmur from that of aortic stenosis. *The American Journal of Cardiology*, 95, 902–904.
- Prasad, D. K., Leung, M. K., & Quek, C. (2012). Ellift: An unconstrained, non-iterative, least squares based geometric ellipse fitting method. *Pattern Recognition*.
- Adolph, R. J., Tanaka, K., & Stephens, J. F. (1970). The clinical value of frequency analysis of the first heart sound in myocardial infarction. *Circulation*, 41, 1003–1014.
- Samjin Choi, J.A.E. (2005). Development of wireless heart sound acquisition system for screening heart valvular disorder. In *Proceedings of the international*

- conference on instrumentation, control and information technology (SICE) (pp. 3771–3776).
- Sava, H., McDonnell, J., & Fox, K. (1994). Spectral analysis of first and second heart sounds before and after mechanical heart valve implantation. In *Engineering in medicine and biology society. Engineering advances: New opportunities for biomedical engineers. Proceedings of the 16th annual international conference of the IEEE* (Vol. 2, pp. 1280–1281). IEEE.
- Shuping Sun, H.W.Y.F.T.T., & Zhongwei Jiang (2013). A novel automatic method for heart sounds segmentation and peak location. *Expert Systems with Applications*. [unpublished results].
- Syed, Z., Leeds, D., Curtis, D., & Guttag, J. (2006). Audio-visual tools for computer-assisted diagnosis of cardiac disorders. In *2012 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS)* (pp. 207–212).
- Reed, T. R., Fritzsos, P., & Reed, N. E. (2004). Heart sound analysis for symptom detection and computer-aided diagnosis. *Simulation Modelling Practice and Theory*, 12, 129–146.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York, NY, USA: Springer-Verlag, Inc.
- Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Trans. Neural Netw.*, 10, 988–999.
- VSDCause2013 (2013). Children may be underweight, but may not have cardiac failure. <<http://answers.yahoo.com/question/index?qid=2009052500-4800AApDuvG>>.
- Wikipedia (2013). Ventricular septal defect (vsd). <[http://en.wikipedia.org/wiki/Ventricular\\_septal\\_defect](http://en.wikipedia.org/wiki/Ventricular_septal_defect)>.
- Wu, C. -H., Lo, C. -W., & Wang, J. -F. (1995). Computer-aided analysis and classification of heart sounds based on neural networks and time analysis. In *1995 International conference on acoustics, speech, and signal processing, ICASSP-95, Vol. 5* (pp. 3455 –3458).
- Yadollahi, A., & Moussavi, Z. M. (2006). A robust method for heart sounds localization using lung sounds entropy. *IEEE Transactions on Biomedical Engineering*, 53, 479–502.
- Yan, Z., Jiang, Z., Miyamoto, A., & Wei, Y. (2010). The moment segmentation analysis of heart sound pattern. *Computer Methods and Programs in Biomedicine*, 98, 140–150.