

On the Mahalanobis Distance Classification Criterion for a Ventricular Septal Defect Diagnostic System

Shuping Sun, Xiaohang Li, Hao Chen, Biqiang Zhang, Tingting Huang, Hongxiang Pang

Nanyang Institute of Technology, Nanyang Henan
Email: shp_sun@yeah.net Xiaohang_li_top@163.com

Received: Feb. 27th, 2020; accepted: Mar. 13th, 2020; published: Mar. 20th, 2020

Abstract

This study proposes a criterion based on the Mahalanobis distance for diagnosing three-type ventricular septal defects (VSDs). The three stages corresponding to the diagnostic method are generally summarized as follows. In the first stage, the heart sound is collected via an electronic stethoscope and preprocessed using the wavelet decomposition. The time-frequency features are extracted in the second stage. And finally, the third stage describes the Mahalanobis distance classification criterion-based diagnostic method used to diagnose the VSD. The performance of this proposed method is evaluated by comparing with other well-known classification methods in diagnosing sounds from patients with VSDs. The classification accuracy of three-kind of VSDs and normal heart sound are 95.2%, 94.4%, 97.1%, and 99.1%, respectively, which are greater than other well-known classifier methods. Therefore, the proposed method can provide an efficient way to diagnose VSD for medical staff or patients.

Keywords

Mahalanobis Distance, Chi-Square Distribution, Gaussian Mixture Models, Ventricular Septal Defect, Principle Component Analysis

基于马氏距离分类准则的室间隔缺损诊断研究

孙树平, 李肖航, 陈豪, 张弼强, 黄婷婷, 庞宏祥

南阳理工学院, 河南 南阳
Email: shp_sun@yeah.net Xiaohang_li_top@163.com

收稿日期: 2020年2月27日; 录用日期: 2020年3月13日; 发布日期: 2020年3月20日

摘要

本研究提出一种基于马氏距离的分类准则实现对三种室间隔缺损(VSD)的诊断。与诊断方法对应的三个阶段概括如下。第一阶段,通过电子听诊器采集心音,然后利用小波分解进行预处理。在第二阶段提取时频特征并进行主成分分析(PCA)降维。第三阶段,描述了基于马氏距离分类准则的VSD诊断方法。最后,通过与其它诊断VSD的分类方法进行比较来评价本方法效果。分析研究结果,本研究对三种室间隔缺损与正常心音的分类精度分别为95.2%、94.4%、97.1%、99.1%,优于其它知名分类器。因此本研究可为医护人员或患者诊断VSD提供一种有效的方法。

关键词

马氏距离, 卡方分布, 高斯混合模型, 室间隔缺损, 主成分分析

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

先心病是患病率最高的先天性缺损疾病,而室间隔缺损(VSD)是最常见的先心病类型。据研究[1]统计,室间隔缺损约占先心病的35.6%,全世界每1000名新生儿中约有3名婴儿患有室间隔缺损。如图1所示,为全球患病率最高的五种先心病病种在1970~2017年患病率变化趋势图。在临床诊断中,室间隔缺损通常根据缺损直径的大小分三种类型[2]:小型室间隔缺损(SVSD)、中型室间隔缺损(MVSD)、大型室间隔缺损(LVSD)。其中,SVSD的直径小于等于5 mm, MVSD的直径大于5 mm并小于等于15 mm, LVSD的直径大于15 mm。诊断室间隔缺损的常用方法[3]一般有胸部X-射线透视、心电图、临床听诊等。研究[3][4]表明,最简便有效的办法是心音听诊,并且从早期诊断开始一直沿用至今。然而单一地听诊对正常心音和室间隔缺损心音的分辨精度仅达90%,室间隔缺损的缺损直径大小将直接影响后期的诊断和理疗,因此,提高室间隔缺损诊断的分类精度显得尤为重要。鉴于此国内外大量研究[2][5]提出了多种心音识别算法,被广泛应用于区分心脏疾病和评估心脏功能。

而近几年来,在模式识别领域中,高斯混合模型(GMMs)作为一种常用的模式识别方法,被广泛应用于分类和聚类分析。有大量研究[6][7][8][9]由于对高斯函数特征的误解或错误地与一维高斯函数类比,大量使用二维高斯模型的研究未指定或未能建立正确的置信区间,导致分类精度不高。为解决这种问题,本研究提出一种基于马氏距离分类准则的分类算法,寻求马氏距离与高斯混合模型之间的关系,建立最佳的置信区间及其对应的置信水平,并应用于室间隔缺损诊断,实现对正常心音信号和室间隔缺损心音信号更高精度的识别与分类。

2. 方法

本研究提出一种基于马氏距离分类准则的室间隔缺损诊断研究,能够实现对三种室间隔缺损(VSDs)心音信号进行分类诊断。最后通过分类精度,灵敏度,特异性值来验证所提方法的有效性。其算法总体流程图如图2所示。

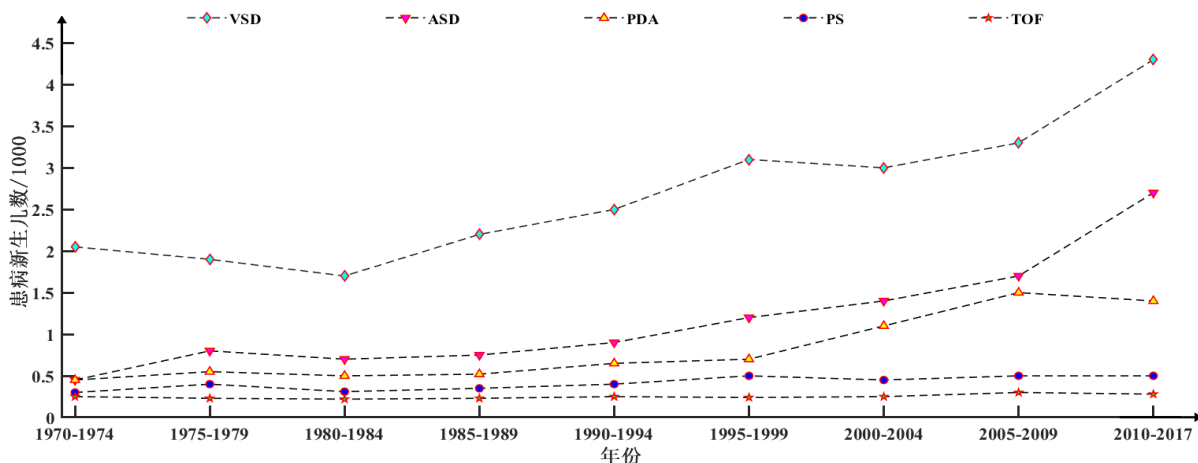


Figure 1. Trend chart of global prevalence of five congenital heart diseases from 1970 to 2017

图 1. 1970~2017 五种先天性心脏病全球患病率变化趋势图

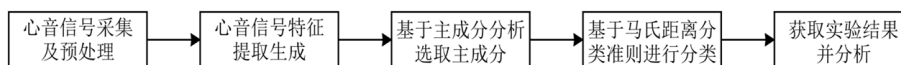


Figure 2. Flow chart of Mahalanobis distance classification criterion algorithm

图 2. 马氏距离分类准则算法流程图

2.1. 心音信号的采集及预处理

在临床医学上, 心音听诊区[10]主要分为四个, 分布在心脏瓣膜位置, 如图 3 所示。其中 A、P、M、T 分别为主动脉瓣听诊区, 肺动脉瓣听诊区, 二尖瓣听诊区, 三尖瓣听诊区。3M®公司的 3M-3200 型电子听诊器不仅可以实时存储心音信号, 而且能够导出 wav 文件, 便于进行后续分析处理, 所以本研究采用 3M-3200 型电子听诊器采集记录心音。采样频率设置为 44.1 kHz, 灵敏度调节至最高档, 频带宽度设置为 20 Hz~20 kHz, 记录时间 30~60 s。本研究所收集的正常心音来自南阳理工学院的健康个体, 室间隔缺损患者心音信号来自于 3M 公司数据库[11]。心音信号在采集过程中会受到各种因素的影响及产生的噪声干扰, 导致部分信息量损失, 心音质量也大幅下降, 同时将对后续的心音信号分析处理产生一定的误差。鉴于心音信号所具有的复杂性和不稳定性以及引入的噪声增加了心音分析的难度, 因此采集过程中的降噪处理与心音分析中的去噪处理显得尤为重要。小波分解被广泛应用于心音信号的预处理, 根据文献[12]-[18]对心音频率范围的研究, 利用小波分解对心音信号进行滤波处理, 保留大于 21.5 Hz, 小于 689 Hz 的频率成分。研究[17]通过对 Daubechies 小波系列(Db2-Db10 小波)、Haar 小波、Symlets 小波(Sym2-Sym6)、Cioflet 小波系列(Ciof1-Ciof5)和 BiorSplines 小波系列(Bior1.1-Bior3.3)进行测试, 总结得到 Db10 小波和离散 Meyer 小波能够使心音获得最大信噪比和最小均方差。因此, 本研究采用 Db10 作为母小波对所采集的心音信号进行五层分解, 结果如图 4 所示。图 4(a)为带噪的心音信号, 图 4(b)为经过 Db10 小波五层分解降噪后的波形。通过心音波形可以看出, 采用 Db10 小波对心音进行小波分解, 可以使得高频噪声被很好地滤除, 且合理的低频成分被很好地保留, 极大地减少有用信息量的缺失, 获得较好的去噪效果。

2.2. 心音信号的特征提取与特征生成

提取心音信号特征是采用心音信号数据的统计学参数, 根据研究[2]表明, 由时域特征和频域特征组成的时频特征能够提供极其有效的信息量。时频特征向量主成分如表 1, 提取结果如图 5 所示, 其中青色线 S 为频域波形, 红色线 E 为频域包络, 黑色虚线 H 代表阈值线(其设定值为 0.2), 绿色三角形 L 和蓝色三角形 R 分别为阈值线与包络的左右交点, 黄色点 f_g 代表频率重心, f_w 为频率宽度。

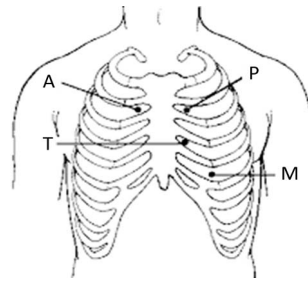


Figure 3. Heart sound auscultation
图 3. 心音听诊区

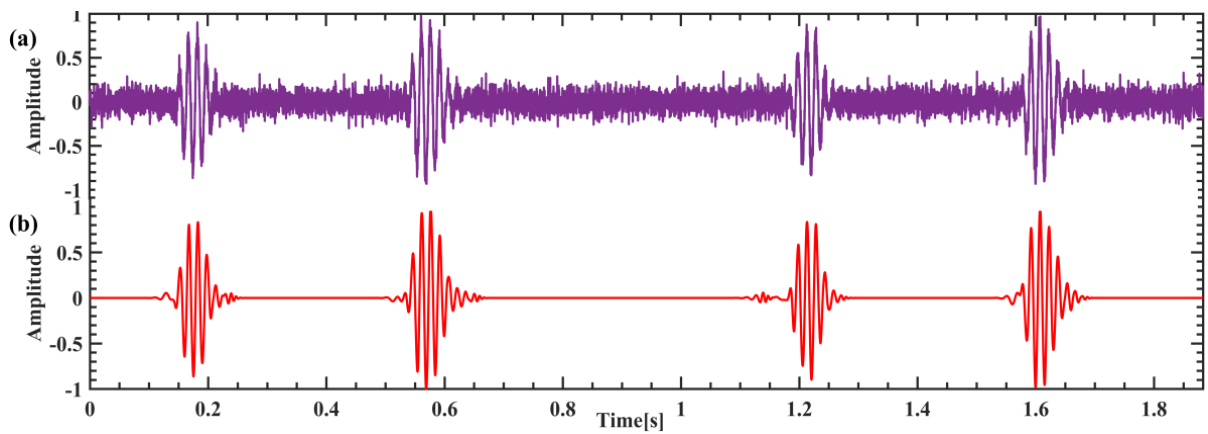


Figure 4. Preprocessing experimental results
图 4. 预处理实验结果图

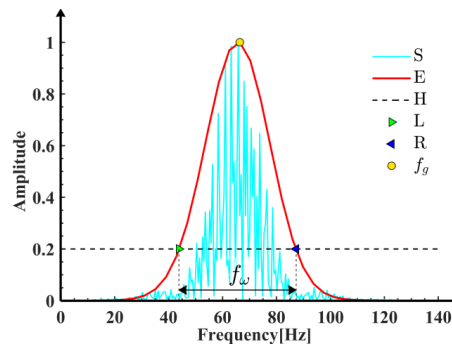


Figure 5. Frequency domain feature extraction result diagram
图 5. 频域特征提取结果图

Table 1. Time frequency characteristic table of principal components
表 1. 主成分时频特征表

Features	Principal components			
	v_1	v_2	v_3	v_4
t_{12}	0.3601	0.6662	0.6551	-0.0312
t_{11}	0.5131	0.4501	-0.7216	-0.0759
f_g	-0.5327	0.4520	-0.1821	0.6901
f_w	-0.5723	0.3836	-1123	-0.7226

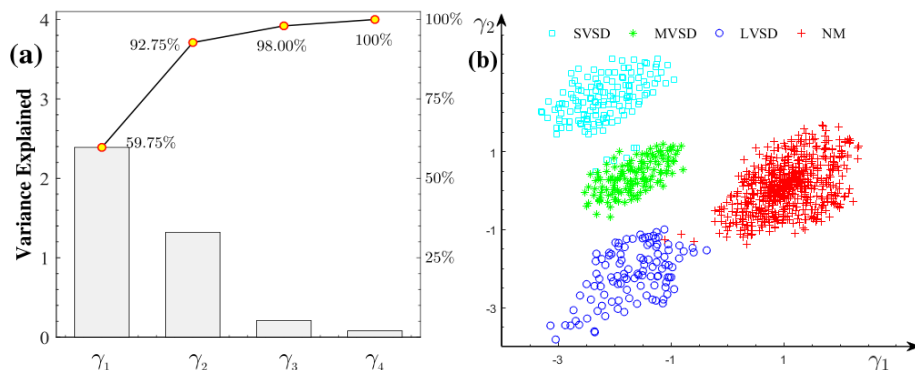


Figure 6. PCA result chart
图 6. 主成分分析结果图

鉴于主成分分析(PCA)是一种线性降维技术,可为高维数据进行降维处理获取新的主成分,在大量研究中被广泛应用于诊断系统。如图 6(a)所示为帕累托图,代表了各个主成分比重,明显看出主成分为 γ_1 和 γ_2 。而图 6(b)为第一主成分和第二主成分的散点图,可明显确定 γ_1 和 γ_2 的分布区间有四个区域,其中青色方框代表 SVSDs,绿色星号代表 MVSDs,蓝色圆圈代表 LVSDs,红色加号代表正常心音(NM)。从图中可以看出,各类心音信号特征之间是具有一定差别区分的,且两个主成分 γ_1 和 γ_2 所占有的信息量占据总体的 92.75%,可以代表整个心音特征,因此,将 γ_1 和 γ_2 作为主成分,将其数据提取作为马氏距离分类准则的分类原始数据,并为进行下一节中的建立高斯混合模型以及成分参数估计做准备。

2.3. 室间隔缺损诊断

作为一种模式识别方法,高斯混合模型(GMMs)在统计学中具有重要作用,被广泛用于分类识别或聚类分析算法[19] [20]。当使用 GMMs 进行决策时,实际上必须指定分类区间对应的置信水平。而大量的研究在使用二维高斯分布时,由于对高斯函数的错误理解或盲目地与一维正态分布进行类比,最终未建立正确的置信区间。鉴于马氏距离的平方服从 χ^2 分布[21],可借助 GMMs 实现置信区间对检测数据进行分类,最终可以确定基于 GMMs 的最佳置信区间的综合推导。我们将此定义为马氏距离分类准则的推导过程。马氏距离分类准则的定义推导分为以下五个步骤:

第一步:寻找估计目标函数 $f(x)$,采用极大似然估计理论获取高斯混合参数。能够作为 n 维标准高斯分布的混合,尽可能产生每个成分的训练模型,并对每个成分进行参数估计。如下公式(1) (2):

$$f(x) = \sum_{k=1}^K \omega_k p(x | \mu_k, \Sigma_k) \tag{1}$$

$$p(x | \mu_k, \Sigma_k) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_k|}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)} \tag{2}$$

其中 $p(x | \mu_k, \Sigma_k)$ 是每个成分的后验概率, K 是成分数量, ω_k 对应于第 k 个混合模型的加权系数且 $\sum_{k=1}^K \omega_k = 1$, μ_k 和 Σ_k 分别是第 k 个成分的样本均值和协方差矩阵。根据图 6(b)所示的散点图,可明显观察到有四个成分。而又根据公式(1) (2)所示,因此设置高斯混合成分数量 $K = 4$,并利用 Matlab2017b 的软件平台,采用 *fitgmdist* 将生成的特征拟合,建立一个 GMM 结构。根据经验设置正则化值为 0.01,优化迭代数目为 1000。利用高斯混合参数估计获得 ω_k 、 μ_k 和 Σ_k ,具体参数数值如表 2 所示。并且对四个成分的高斯混合参数估计所获得的网格图如图 7 所示,其中成分编号 1~4 分别对应 SVSD、MVSD、LVSD 和 NM。由图可知这四个随机成分没有重叠区域,也就表明每个成分之间是相互独立的。而在二维平面上可以观测到对于每个独立成分的马氏距离平方都具有一个椭圆轮廓,服从 χ^2 分布。

Table 2. Gaussian mixture parameter estimation table
表 2. 高斯混合成分参数估计表

Components	Components number	Gaussian mixture parameter estimates				
		w_k	μ_k	Σ_k	μ_k	Σ_k
SVSD Classifier	$k = 1$	0.1213	-2.1707	2.3246	0.2770	0.1244
					0.1244	0.3301
MVSD Classifier	$k = 2$	0.1332	-1.6615	0.3677	0.1841	0.1166
					0.1166	0.1793
LVSD Classifier	$k = 3$	0.1326	-1.6764	-2.1499	0.3230	0.2118
					0.2118	0.4625
NM Classifier	$k = 4$	0.6129	1.0554	0.0838	0.3043	0.1617
					0.1617	0.3875

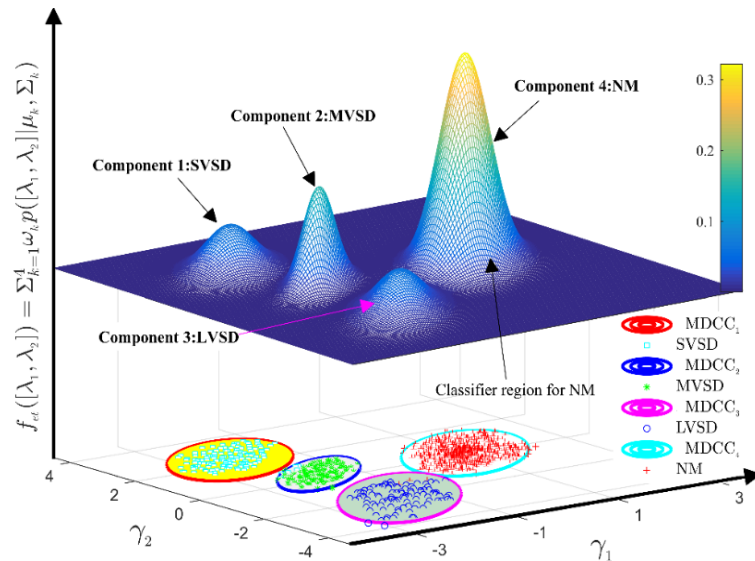


Figure 7. Grid diagram and contour diagram
图 7. 网格图与等高线图

第二步：为获得对待测数据分类的决策区间，通过第一步成分的参数估计，样本点到第 k 个成分的马氏距离的平方 (MD_k^2) 计算公式如公式(3)：

$$MD_k^2 = (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) \quad (3)$$

观察公式(2)与公式(3)，公式(2)满足标准正态分布的形式，所以可得马氏距离的平方服从 χ^2 分布如下：

$$MD_k^2 \sim \chi^2(\mu_k, \Sigma_k) \quad (4)$$

室间隔缺损诊断的分类区间，可以通过指定需要的置信水平 α_k ，利用 χ^2 分布特性来获得。

第三步：确定每个成分的置信水平 (α_k)。通过设定指定的置信水平得到每个成分的置信区间，每个成分的分类区间可由一个椭圆区域所确定，表示为：

$$MD_k^2 \leq MDCC_k = \chi_{1-\alpha_k}^2 \quad (5)$$

其中, $MDCC_k$ 代表的就是每个成分对应的分类准则, 即马氏距离分类准则。并且满足下面的等式(6):

$$P\{\chi^2 = MD_k^2 \leq MDCC_k\} = \alpha_k \tag{6}$$

对于 χ^2 分布, 大量研究[22]-[27]都广泛采用置信水平分别为 68.3%、95%、97.5%作为分类的准则, 鉴于每个类别都需要特定的置信水平, 必须确定最优置信水平 α_k 。 α_k 的确定是通过设置 $\alpha \in [68.3\% : 0.1\% : 99\%]$ 并依据每个椭圆轮廓尽可能大且公共区域尽可能小的原则, 最后确定结果为: $\alpha_1 = 0.97$, $\alpha_2 = 0.975$, $\alpha_3 = 0.962$ 和 $\alpha_4 = 0.95$ 。

第四步: 通过上一步确定的 α_k , 利用 Matlab 平台关于卡方分布的逆概率分布函数(chi^2inv)计算每个成分的 $MDCC_k$ 。其中 k 从 1 到 4 分别对应 SVSD、MVSD、LVSD 和 NM。

第五步: 基于马氏距离分类准则定义室间隔缺损诊断方法如公式(7):

$$\begin{cases} \text{SVSD, } MD_1^2 \leq MDCC_1 \\ \text{MVSD, } MD_2^2 \leq MDCC_2 \\ \text{LVSD, } MD_3^2 \leq MDCC_3 \\ \text{NM, } MD_4^2 \leq MDCC_4 \\ \text{未知类别, 其他} \end{cases} \tag{7}$$

最后, 为评估这些椭圆模型的性能, 其分类精度(CA), 灵敏度(Se)和特异性值(Sp)通过公式(8)计算。其中 TP、FP、TN 和 FN 分别是真阳性、假阳性、真阴性和假阴性的数量。

$$\begin{cases} CA(\%) = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \\ Se(\%) = \frac{TP}{TP + FN} \times 100\% \\ Sp(\%) = \frac{TN}{FP + TN} \times 100\% \end{cases} \tag{8}$$

3. 马氏距离分类准则分类结果及验证

按照 2.3 节步骤在 Matlab2017b 中运行马氏距离分类准则算法, 获得三种室间隔缺损心音(SVSD、MVSD、LVSD)和正常心音(NM)的马氏距离分类准则($MDCC_k$), 分别为: 7.01、7.38、6.54、5.99。同时, 根据置信水平为 $\alpha = 0.95$ 对应的 χ^2 分布置信区间, 可用来诊断室间隔缺损心音与正常心音。诊断结果如图 8 所示。其中四种星号分别代表四种心音信号的样本中心, 其椭圆外围对应四种心音信号的分类区间。经实验验证获得三种室间隔缺损和正常心音的分类精度, 灵敏度和特异性值如下表 3 所示。

Table 3. Ventricular septal defect diagnostic results
表 3. 室间隔缺损诊断结果

类别	MDCC	CA (%)	Se (%)	Sp (%)
SVSD	7.01	95.2	96.7	94.9
MVSD	7.38	94.4	95.1	94.3
LVSD	6.54	97.1	97.7	96.5
NM	5.99	99.1	98.8	99.5

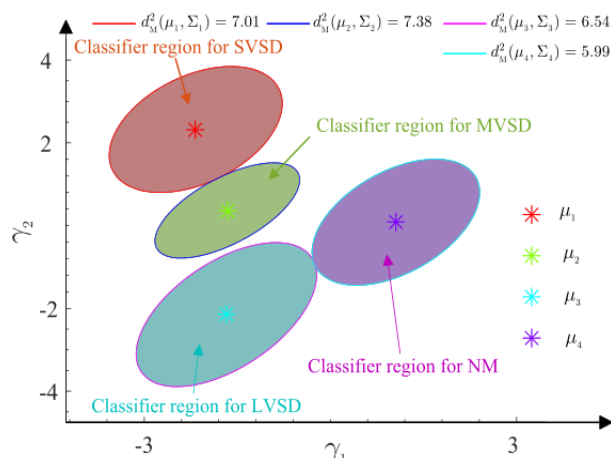


Figure 8. Classification of ventricular septal defects
图 8. 室间隔缺损分类结果图

4. 结论

本研究针对心音特征构建了高斯混合模型的最佳置信区间，用于诊断室间隔缺损(VSD)。为此，提出了马氏距离分类准则与基于 PCA 的心音特征生成相结合的方法。为简化时频特征，通过 PCA 生成前两个主要分量 γ_1 和 γ_2 ，用作诊断特征。首先针对 VSD 和正常心音建立关于 $[\gamma_1, \gamma_2]$ 的高斯混合模型，然后根据卡方分布在置信水平 ($\alpha = 0.95$) 下的置信区间定义了马氏距离分类准则，用于诊断 VSD 和正常心音。本研究所提方法，目前正在尝试应用于其他心脏类疾病。并且通过整个研究过程了解到，目前我国在计算机技术辅助医疗诊断方面的技术还是比较缺乏的。因此，需要我们对这方面有更加深入的研究，健全我国健康医疗辅助技术知识体系。

参考文献

- [1] Liu, Y., et al. (2019) Early Childhood Global Birth Prevalence of Congenital Heart Defects 1970-2017: Updated Systematic Review and Meta-Analysis of 260 Studies. *International Journal of Epidemiology*, **48**, 455-463. <https://doi.org/10.1093/ije/dyz009>
- [2] Sun, S., Wang, H., Jiang, Z., Fang, Y. and Tao, T. (2014) Segmentation-Based Heart Sound Feature Extraction Combined with Classifier Models for a VSD Diagnosis System. *Expert Systems with Applications*, **41**, 1769-1780. <https://doi.org/10.1016/j.eswa.2013.08.076>
- [3] Bhatikar, S.R., DeGroff, C. and 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. <https://doi.org/10.1016/j.artmed.2004.07.008>
- [4] Higuchi, K., Sato, K., Makuuchi, H., Furuse, A., Takamoto, S. and Takeda, H. (2006) Automated Diagnosis of Heart Disease in Patients with Heart Murmurs: Application of a Neural Network Technique. *Journal of Medical Engineering & Technology*, **30**, 61-68. <https://doi.org/10.1080/03091900500131110>
- [5] 谭朝文, 王威廉, 宗容, 等. 卷积神经网络应用于先心病心音信号分类研究[J]. *计算机工程与应用*, 2019, 55(12): 174-180.
- [6] Nieto, M., Cuevas, C. and Salgado, L. (2009) Measurement-Based Reclustering for Multiple Object Tracking with Particle Filters. *International Conference on Image Processing*, Cairo, 7-10 November 2009, 4097-4100. <https://doi.org/10.1109/ICIP.2009.5413709>
- [7] Suhr, J.K., Jung, H.G., Li, G. and Kim, J. (2011) Mixture of Gaussians-Based Background Subtraction for Bayer-Pattern Image Sequences. *IEEE Transactions on Circuits and Systems for Video Technology*, **21**, 365-370. <https://doi.org/10.1109/TCSVT.2010.2087810>
- [8] Hassanpour, H., Sedighi, M. and Manashty, A.R. (2011) Video Frame's Background Modeling: Reviewing the Techniques. *Journal of Signal and Information Processing*, **2**, 72-78. <https://doi.org/10.4236/jsip.2011.22010>

- [9] Gallego, J., Pardàs, M. and Haro, G. (2012) Enhanced Foreground Segmentation and Tracking Combining Bayesian Background, Shadow and Foreground Modeling. *Pattern Recognition Letters*, **33**, 1558-1568. <https://doi.org/10.1016/j.patrec.2012.05.004>
- [10] Color, S. and Cues, V. (2013) Heart Sounds. iWorx Physiology Lab Experiment, 03820, 1-8.
- [11] Littmann Library, Ventricular Septal Defect Database, 3M Company. <http://www.3m.com/healthcare/littmann/pn111.html>
- [12] 高巍, 彭宇. 基于马氏距离多核学习的高光谱图像分类[J]. 仪器仪表学报, 2018(3): 250-257.
- [13] Nilsson, M., Funk, P., Olsson, E.M.G., Von Schéele, B. and 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. <https://doi.org/10.1016/j.artmed.2005.04.004>
- [14] Jabbari, S. and Ghassemian, H. (2009) Sparse Modeling of Heart Sounds and Murmurs Based on Orthogonal Matching Pursuit. *14th CSICC*, Tehran, 20-21 October 2009, 355-360. <https://doi.org/10.1109/CSICC.2009.5349606>
- [15] Choi, S. and 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. <https://doi.org/10.1016/j.compbiomed.2009.10.003>
- [16] Choi, S., Shin, Y. and Park, H.K. (2011) Selection of Wavelet Packet Measures for Insufficiency Murmur Identification. *Expert Systems with Applications*, **38**, 4264-4271. <https://doi.org/10.1016/j.eswa.2010.09.094>
- [17] Ali, M.N., El-Dahshan, E.S.A. and Yahia, A.H. (2017) Denoising of Heart Sound Signals Using Discrete Wavelet Transform. *Circuits, Systems, and Signal Processing*, **36**, 4482-4497. <https://doi.org/10.1007/s00034-017-0524-7>
- [18] Hamidi, M., Ghassemian, H. and Imani, M. (2018) Classification of Heart Sound Signal Using Curve Fitting and Fractal Dimension. *Biomedical Signal Processing and Control*, **39**, 351-359. <https://doi.org/10.1016/j.bspc.2017.08.002>
- [19] 肖应旺, 杨军, 张承忠, 等. 统计监控建模数据预处理离群点检测算法[J]. 控制工程, 2013, 20(4): 756-761.
- [20] De Maesschalck, R. and Massart, D.L. (2000) The Mahalanobis Distance. *Chemometrics and Intelligent Laboratory Systems*, **50**, 1-18. [https://doi.org/10.1016/S0169-7439\(99\)00047-7](https://doi.org/10.1016/S0169-7439(99)00047-7)
- [21] Etherington, T.R. (2019) Mahalanobis Distances and Ecological Niche Modelling: Correcting a Chi-Squared Probability Error. *PeerJ*, **7**, e6678. <https://doi.org/10.7717/peerj.6678>
- [22] Gallego, G., Cuevas, C., Mohedano, R. and García, N. (2013) On the Mahalanobis Distance Classification Criterion for Multidimensional Normal Distributions. *The IEEE Transactions on Signal Processing*, **61**, 4387-4396. <https://doi.org/10.1109/TSP.2013.2269047>
- [23] Qiu, L., Yuan, S., Chang, F.K., Bao, Q. and Mei, H. (2014) On-Line Updating Gaussian Mixture Model for Aircraft Wing Spar Damage Evaluation under Time-Varying Boundary Condition. *Smart Materials and Structures*, **23**, Article ID: 125001. <https://doi.org/10.1088/0964-1726/23/12/125001>
- [24] Pinto, R.C. and Engel, P.M. (2015) A Fast Incremental Gaussian Mixture Model. *PLoS ONE*, **10**, e0139931. <https://doi.org/10.1371/journal.pone.0141942>
- [25] Proïa, F., Pernet, A., Thouroude, T., Michel, G. and Clotault, J. (2016) On the Characterization of Flowering Curves Using Gaussian Mixture Models. *Journal of Theoretical Biology*, **402**, 75-88. <https://doi.org/10.1016/j.jtbi.2016.04.022>
- [26] Mungai, P.K. and Huang, R. (2017) Using Keystroke Dynamics in a Multi-Level Architecture to Protect Online Examinations from Impersonation. *IEEE 2nd International Conference on Big Data Analysis*, Beijing, 10-12 March 2017, 622-627. <https://doi.org/10.1109/ICBDA.2017.8078710>
- [27] Aryafar, A., Mikaeil, R., Doulati Ardejani, F., Shaffiee Haghshenas, S. and Jafarpour, A. (2018) Application of Non-Linear Regression and Soft Computing Techniques for Modeling Process of Pollutant Adsorption from Industrial Wastewaters. *Journal of Mining and Environment*, **10**, 327-337.