

# 有效特征参数分类正常与病理语音

郭乐乐, 曹辉, 李涛

(陕西师范大学物理学与信息技术学院, 陕西西安 710100)

**摘要:** 采用残差信号的特征参数——基音幅值(Pitch Amplitude, PA)和频谱平坦度(Spectral Flatness of the Residue Signal, SFR)与语音信号倒谱域特征参数——倒谱峰值突出(Cepstral Peak Prominence, CPP)来区分正常与病理语音, 在萨尔布吕肯语音数据库中选择自然音调的正常与病理语音/a/进行仿真实验。统计结果表明, 与正常语音相比, 病理语音的 PA 较小, SFR 更接近零, CPP 也较小。结合其他传统特征参数分析对比, 证明 SFR、PA 和 CPP 更能有效分类正常与病理语音。通过不同分类算法比较, 得出支持向量机的分类准确率相对更高。

**关键词:** 残差信号; 基音幅值; 频谱平坦度; 倒谱峰值突出; 支持向量机

中图分类号: H107

文献标识码: A

文章编号: 1000-3630(2019)-05-0554-06

DOI 编码: 10.16300/j.cnki.1000-3630.2019.05.012

## Classification of normal and pathological speech by effective feature parameters

GUO Le-le, CAO Hui, LI Tao

(School of Physics and Information Technology, Shaanxi Normal University, Xi'an 710100, Shaanxi, China)

**Abstract:** The feature parameters PA (pitch amplitude) and SFR (spectral flatness of the residue signal) and the vowel cepstrum domain feature parameter CPP (cepstral peak prominence) are used to distinguish between normal and pathological speeches. In the Saarbruecken Voice Database, 216 normal and 216 pathological natural tones /a/ are selected for experiments. The statistical results show that compared with normal speech, the PA value of pathological speech is smaller, the SFR value is close to zero, and the CPP value is also smaller. Combined with other features analysis and comparison, it is proved that SFR, PA, and CPP are excellent and stable feature parameters for normal and pathological speech classification. The classification accuracy obtained by support vector machine is relatively higher by the comparison of different classification algorithms.

**Key words:** residue signal; pitch amplitude (PA); spectral flatness of the residue signal (SFR); cepstral peak prominence (CPP); support vector machine

## 0 引言

导致语音异常的因素有很多种, 其中不同声带病理类型造成的嗓音疾病是语音异常最主要的因素。嗓音疾病影响声带的形态和振动模式, 使得发声过程中声带异常闭合或者左右不对称振动, 造成嗓音的声学特征和空气动力学特征等均发生改变, 导致嗓音嘶哑、无力等, 严重影响了人们的生活质量和社会交流<sup>[1]</sup>。常见的嗓音疾病主要有声带麻痹、声带水肿、声带息肉、声带小结等, 这些嗓音疾病患者发出的声音统称为病理语音。所有病理语音的形成大多和声带振动异常有关, 这导致病理语音中

包含更多无声或不规则有声成分, 病理严重时语音波形特征类似噪声。

病理嗓音检测难度与病理的严重程度密切相关, 病理语音的特征选择在病理语音识别中具有重要地位。2014 年, 张满彩<sup>[2]</sup>直接从原始语音中提取传统声学特征对病理声音进行了分析, 结果表明传统声学特征不能很好地表征语音信号的动态特征问题。除了直接从语音信号中提取特征分析, 人们还常用声门波相关特征参数来量化语音, 2016 年, BURROWS 等<sup>[3]</sup>提取了声门波的时域和频域特征参数用来辨别声带小结、声带麻痹与正常说话者, 结果表明与传统声学特征相比, 声门波特征的分类效果更好。之后 FORERO 等<sup>[4]</sup>又将声门波特征与梅尔倒谱系数结合在一起, 对正常语音与病理语音进行分类, 分类率得到了提高。

常见的提取声门波的方法主要有两种, 分别为最小二乘法和基音迭代自适应法, 二者都是基于线

收稿日期: 2018-05-07; 修回日期: 2018-07-18

基金项目: 国家自然科学基金资助项目(11074159、11374199)

作者简介: 郭乐乐(1994—), 女, 山西临汾人, 硕士研究生, 研究方向为信号与信息处理。

通讯作者: 曹辉, E-mail: caohui@snnu.edu.cn

性预测 (Linear Prediction, LP) 逆滤波实现的。但是最小二乘法对被测语音要求较高, 只有当被测语音具有足够长的闭合相且这些闭合相位置被准确测出时, 求得的声门波才可靠<sup>[5]</sup>。大部分病理说话人发声时声带不能正常闭合, 发出长时闭合相的语音相对困难, 因此, 最小二乘法不太适合病理语音声门波的提取, 局限性较大。考虑到实际操作, 一般采用基音迭代自适应法提取正常与病理语音声门波<sup>[6]</sup>。首先通过一阶 LP 分析初步估计声门波对语音信号的影响, 然后逆滤波消除这种影响, 再通过更高阶的 LP 分析得到声道的首次估计, 最后进行逆滤波和积分依次消除声道和唇辐射的影响, 至此完成第一次迭代过程, 得到初始声门波。将第一次迭代的输出作为第二次迭代的输入, 重复第一次迭代过程, 得到最终的声门波。该方法对声门波和声道传输函数重复两次迭代, 有效避开了闭合相测量, 鲁棒性高, 但需要对语音信号进行多次线性预测, 过程相对复杂。

目前定量分析声门波使用的时域特征参数有开放商(开放相与基音周期的比值)、闭合商(闭合相与基音周期的比值)和速度商(渐开相与渐闭相的比值)等, 这些参数需要准确判定声门的开放和闭合时刻, 而声带逐渐开放或不完全闭合等会导致这些关键时间点求解出错, 共振峰波纹和噪声也会使量化的准确度下降。声门波的频域特征参数谐波差(一次谐波与二次谐波的幅值差)和谐波丰富因子(其它谐波的幅值总和与一次谐波幅值的比值)的求解虽不需要确定具体时间点, 但它们不能有效捕捉时间细节的快速变化, 所以合理量化声门波也比较困难<sup>[5]</sup>。除此之外, 逆滤波得到的声门波本身主要包含低频信息, 因而用此类特征参数来判断正常语音与病理语音相对不易<sup>[7]</sup>。故本文主要从分析嗓音障碍的残差信号入手, 求其特征参数对正常语音与病理语音进行分类。通过线性预测分析得到残差信号来量化语音, 分析残差信号的相关特征, 用统计学方法对正常语音与病理语音的残差特征进行比较分析, 并采用支持向量机二分类处理, 充分证明残差信号特征判别正常语音与病理语音的有效性和优越性。

## 1 残差信号

利用语音信号过去若干个采样值的线性组合预测或估计其当前采样值, 可以得到语音的线性预测值, 实际值与预测值之差称为残差信号。

图 1(a)、1(b)分别为正常语音和病理语音的一

帧信号及其对应的残差信号的波形图。正常语音的残差信号在声门闭合处通常包含明显尖峰, 两尖峰之间的距离代表一个基音周期长度, 病理语音由于声带异常振动或声门闭合不全导致残差信号的周期性不太明显, 波形杂乱无规律。

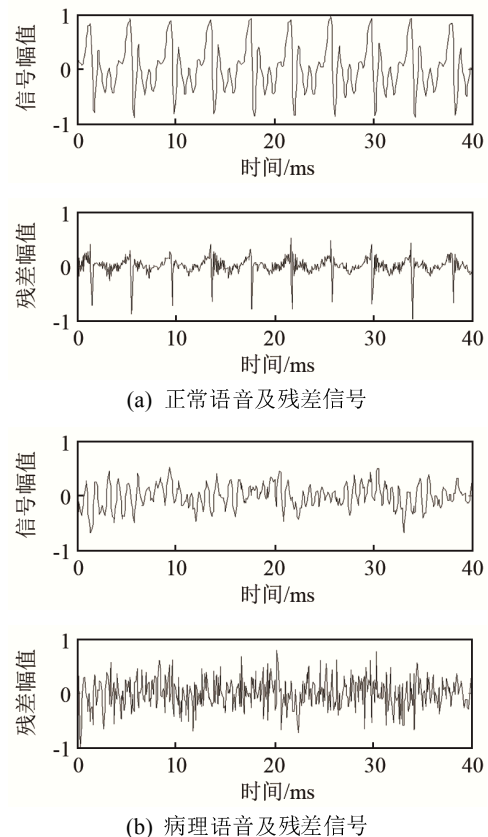


图 1 正常与病理语音及对应的残差信号波形图  
Fig.1 The waveform diagrams of normal and pathological speeches and corresponding residue signals

残差信号提取简单, 只需进行一次线性预测分析就可得到, 它比声门波包含更多的高频信息, 消除了声道的影响, 能够提供声带振动的定时信息。其特征参数的量化也比较客观, 不需要实验者主观干预和运用经验知识。因此, 用它来判断正常语音与病理语音更加合适。

## 2 特征参数

从残差信号中提取出两个特征参数, 分别是基音幅值(Pitch Amplitude, PA)和频谱平坦度(spectral flatness of the residue signal, SFR)<sup>[8]</sup>, 二者统称为残差特征。另外还提取了语音的倒谱域特征参数倒谱峰值突出(Cepstral Peak Prominence, CPP)<sup>[9]</sup>, 用这 3 个特征参数来分析语音, 诊断语音的正常与病理。

### 2.1 基音幅值 PA

基音幅值 PA 是指残差信号求归一化自相关

后, 出现在初始位置之后的最大幅值, 通常对应于波形图中的第二个峰值, 第一个峰值是出现在初始位置的单位幅值 1。语音信号的自相关函数在基频  $F_0$  的周期倍数处出现峰值, 经常用这个特征来检测信号的周期性, 故 PA 可以用来表示信号周期性的强度, 信号周期性越强, PA 值越大, PA 的理论范围在 0~1 之间。

图 2 描述的是图 1 中正常语音和病理语音残差信号的归一化自相关波形。从图 2 中可看出, 正常语音的残差信号自相关最大值很容易辨别, 测得的 PA 值为 0.758 7, 病理语音的残差信号周期性差, 自相关最大值被噪声淹没, 测得的 PA 值为 0.134 5, 明显小于正常语音的 PA 值。

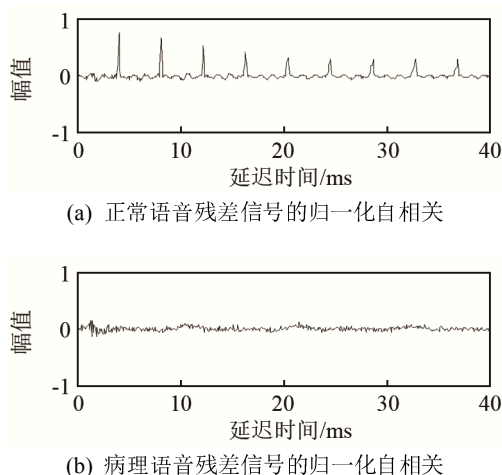


图 2 正常语音与病理语音残差信号归一化自相关波形  
Fig.2 The normalized autocorrelation waveforms of the residue signals of normal and pathological speeches

## 2.2 频谱平坦度 SFR

语音信号频谱主要由基频和谐波组成, 这些频率的分布提供了频谱平坦度 SFR, 它定义为残差信号幅度谱的几何平均值与算术平均值之比的对数, 单位为分贝(dB)。由于几何平均值  $\leq$  算术平均值, 所以 SFR 的理论范围在负无穷与 0 之间。当信号嘈杂, 类似噪声类型分布时, 频谱趋于平坦, 幅度谱的几何平均与算术平均值相差不大, 则 SFR 接近 0; 当频谱由明显的波峰和波谷构成时, SFR 远离 0。喉部病变会造成声带异常振动, 声门处湍流气流增加, 因此影响 SFR 变大(接近 0)。因为频谱是对称的, 所以只显示其频谱的一半, 如图 3 所示。正常语音残差信号的幅度谱中谐波分量丰富明显, 得到的 SFR 为 -6.840 6 dB, 而病理语音残差信号的幅度谱中谐波结构受影响恶化, 不能轻易判断出谐波成分, 得到的 SFR 为 -2.103 4 dB, 比正常残差信号的 SFR 值大。

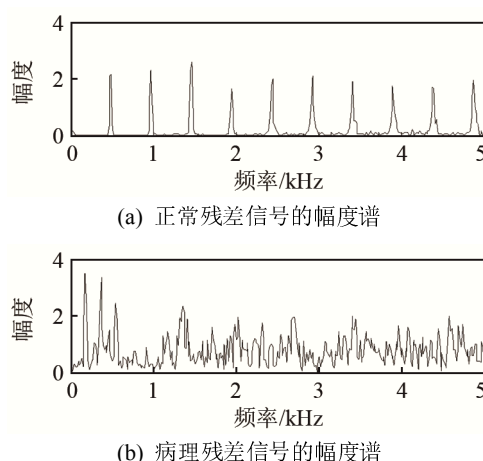


图 3 正常与病理语音残差信号幅度谱  
Fig.3 The amplitude spectra of the residue signals of normal and pathological speeches

## 2.3 倒谱峰值突出 CPP

正常说话人的语音信号具有良好的周期性, 在频域上可以观察到信号有较强的谐波结构; 在倒谱域上则表现为谐波规律性相关的能量波动<sup>[10]</sup>。所谓的倒谱, 是对信号傅里叶变换取对数运算后再进行傅里叶反变换得到的, 是对信号进行两次分析的结果, 相当于对信号求对数功率谱后再求一次对数功率谱, 第一次求的功率谱显示了信号能量的频率分布, 第二次的功率谱说明了频谱中谐波分量的周期性。倒谱的峰值和基频有关, 具有良好的规律性谐波结构的语音信号的倒谱峰值会比较突出, 由此得到特征倒谱峰值突出 CPP, 它定义为倒谱上第一个峰值的幅度和其对应线性回归线上的幅度之差, 单位为 dB。它代表倒谱峰值与倒谱背景噪声之间的距离, 距离越大, 说明离背景噪声越远, 声音信号越清晰。CPP 同样可以反映语音信号的谐波结构水平和噪声分量, 信号中噪声成分越多, 谐波结构越差, CPP 值越小。正常语音周期性成分较多, 发音时平缓, CPP 值较大, 而病理语音由于喉部病变导致语音中噪声成分增加, 谐波结构恶化, 故其 CPP 值下降。

正常语音和病理语音信号的倒谱峰值如图 4 所示, 图 4(c)、4(d)是图 4(a)、4(b)平滑后的结果, 平滑后可以更清楚地观察到倒谱峰值。从图 4(c)、4(d)可知, 正常语音信号在基频周期处会有一个突出的峰值, 而病理语音的周期性较弱, 倒谱峰值不够突出。未平滑之前测得该正常语音的 CPP 值是 34.744 6 dB, 病理语音 CPP 值是 17.630 8 dB, 平滑后正常语音 CPP 值是 20.563 9 dB, 而病理语音 CPP 值为 8.418 2 dB, 可以看出无论是平滑前还是平滑后, 病理语音的 CPP 值都小于正常语音信号的 CPP 值。

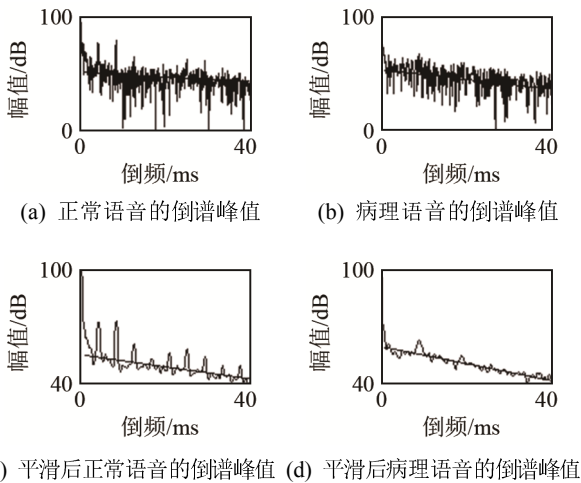


图4 语音信号倒谱峰值图

Fig.4 The cepstral peak diagrams of speech signals

### 3 特征参数仿真实验

#### 3.1 实验预设条件

实验数据来自德国的萨尔布吕肯语音数据库(Saarbruecken Voice Database),数据库中有健康和病理说话者共超过2000人,分别以自然音调、高音调、低音调产生三元音/i/、/a/、/u/的录音文件,文件时长为1~3s,所有记录都是以50kHz采样频率进行采样,分辨率为16bit<sup>[11]</sup>。其语音信号和电声门图(Electroglottography, EGG)信号存储在独立的WAV文件中,任何关于录音的评论都包含在相关的文本文件中。

因为/a/是低元音,发/a/音时喉位低而自然,有利于降低声带和声道之间的耦合作用,可以避开由于鼻音导致逆滤波不精确的问题,故本文实验随机选取了正常和病理共432个人,以自然音调发元音/a/的录音文件,其中正常录音和病理录音各216例。对选好的声音样本进行声学分析,提取所需的声学特征参数,然后进行正常语音与病理语音二分类的研究。

计算残差特征开始前,首先把语音信号采样频率降到11kHz,然后对语音信号进行分帧处理,帧长为1600个采样点,这个长度可以提供正常语音和病理语音的最佳分离<sup>[8]</sup>。然后通过一个14阶线性预测分析滤波器得到每一帧的残差信号,再分别做自相关分析得到时域特征参数PA,做傅里叶变换得到频域特征参数SFR。计算出每一帧的PA和SFR后再求其平均值以得到整个语音的残差特征PA和SFR,作为最终的研究参数。而CPP是将原始语音信号分帧加窗处理后,提取倒谱峰值求得。

#### 3.2 正常语音与病理语音的特征参数比较

采用这432个录音文件分别计算正常语音和病理语音的PA、SFR、CPP参数,求出它们的最大值、最小值、平均值等,得到的实验数据用盒图表示。

盒图主要由五个数值点组成,从下到上依次为:最小值、下四分位数、中位数、上四分位数、最大值。其中下四分位数和上四分位数组成一个空盒子,中位数把盒子一分为二。上四分位数到最大值之间和下四分位数到最小值之间各用一条延伸线连接。盒子中的“×”代表数据平均值,图中的“+”代表离群点,离群点单独绘出的目的是保证整体特征的稳定性,数据不会因此偏移,而盒图中延伸线的两级修改成最小(大)观测值,最小(大)观测值按经验设置为下(上)四分位数减(加)1.5倍四分位数间距离。分析数据时盒图不仅能够帮助我们直观地识别样本数据中的异常值(离群点),而且通过观察盒子的长度、上下间隔的形状以及延伸线的长度可以有效判断样本数据的离散程度和偏向。

为了与其他传统特征参数比较,采用相同的录音文件分别计算元音的谐噪比(Harmonic to Noise Ratio, HNR)、基频微扰Jitter、振幅微扰Shimmer参数。其中,HNR代表信号中谐波能量与噪声能量之比,用来量化声门噪声,可以反映声音质量。信号谐波能量强,HNR大;噪声增加,HNR降低。Jitter指信号相邻周期之间基频的微小变化,Shimmer指相邻周期之间信号幅度的变化,主要反映声带振动的不规律性。发声越稳定,发声过程中声音信号出现的变化越小,Jitter和Shimmer值越小;当声带出现病变后,发声波动程度变大,发音稳定性下降,Jitter和Shimmer值增大,HNR值减小。得到的实验数据如图5所示。

从图5可以看出,正常与病理语音的特征参数之间存在明显差异,由于声带疾病导致病理说话人的声带振动异常,发声不稳定,语音周期性下降,谐波结构恶化,故病理语音的SFR更接近0,PA值更低,CPP值更小;正常语音的基频扰动和振幅扰动较小,信号中谐噪比大,而病理语音的基频扰动和振幅扰动中有更多的离群点,Jitter和Shimmer值增大,HNR值变小。

### 4 支持向量机分类实验

#### 4.1 支持向量机

支持向量机(Support Vector Machine, SVM)是基于统计学习理论的一种机器学习方法,是一种非常有潜力的分类算法,常用于处理二分类问题<sup>[12]</sup>。

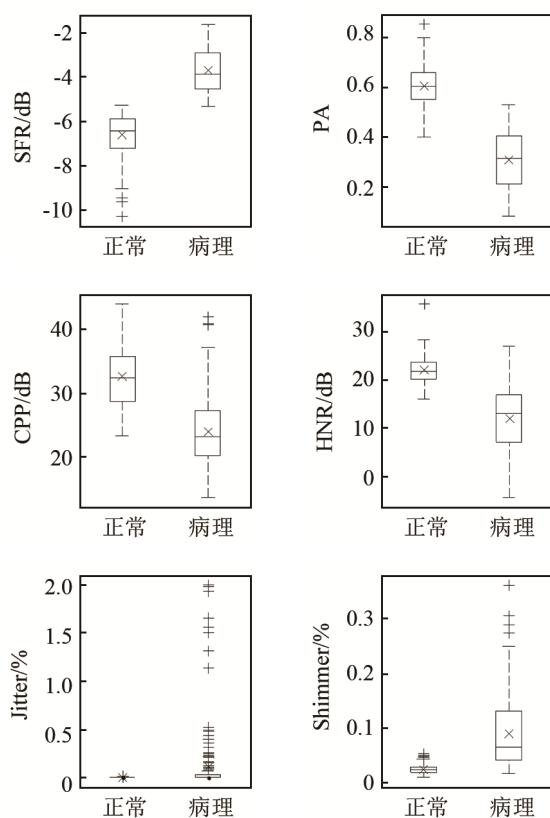


图 5 特征参数的盒图

Fig.5 The box graph of feature parameters

SVM 主要是找到分类超平面把两类样本正确地分开, 即所有相同类别的样本都落在分类超平面的同一侧, 不同类别的样本点之间间隔尽量大, 算法的目的是在无数个分类超平面中选择样本集到分类超平面的距离最大的平面即最优分类面, 也称最大间隔超平面。两类样本中离分类平面最近且平行于最优超平面的分类面上的点就是支持向量, 只要找到所有支持向量即可找到最优超平面, 如果能够找到这个面, 那么这个分类器就称为最大间隔分类器。在进行线性分类时, 将分类面取在离两类样本距离较大的地方; 进行线性不可分时通过高维空间变换转化成高维空间的线性分类问题<sup>[13]</sup>。相比于神经网络算法, 支持向量机采用了核函数, 降低了计算复杂度, 在解决小样本、非线性识别中有其独特的优势。

本文利用支持向量机算法建立正常与病理二分类诊断模型, 经过不同核函数预测正确率对比后, 最终采用默认的径向基(Radial Basis Function, RBF)核函数进行建模。为不失一般性, 训练模型前采用 randperm 函数随机打乱输入的正常和病理语音序列, 设置训练集和测试集之比为 2:1, 故选择语音序列的前 144 个样本作为训练集, 后 72 个样本作为测试集。由于训练集和测试集是随机产生

的, 所以程序每次运行的结果有所不同, 为了客观分析, 取程序运行 10 次得到的平均分类准确率作为最终实验结果。根据测试结果分析评价不同特征参数下诊断模型分类的性能, 分类诊断的大体流程如图 6 所示。

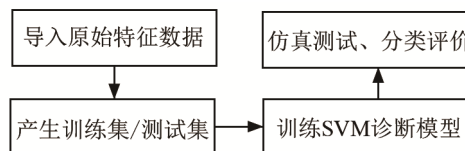


图 6 SVM 二分类诊断模型流程

Fig.6 The flow path of SVM classification diagnosis model

## 4.2 分类识别准确率比较

为了证明 PA、SFR 和 CPP 参数分类识别的优越性, 进一步判断参数的有效性, 客观地区分正常语音与病理语音, 分别对以上 6 个特征用支持向量机进行二分类处理, 得到每个特征参数的分类效果。分类处理后得到的结果如表 1 所示。

分类结果表明, 残差信号特征 PA 和 SFR 的整体分类准确率可以达到 96% 以上, 尤其 SFR 的分类准确率更是高达 99%, 这充分说明了用二者区分正常语音和病理语音的杰出性, 而 CPP 的整体分类准确率高于 90%, 同样可以较好地分类正常语音与病理语音, 但效果比 PA 和 SFR 稍差一些, 对病理语音的分类准确率为 86%, 可能是由于原始语音信号的周期性特征没有残差信号显著, 另外病理语音在时域变化没有在频域变化显著, 这可能同样是 SFR 分类效果较好的原因。

表 1 不同特征参数的分类准确率

Table 1 Classification accuracy of different feature parameters

特征参数	分类准确率/%		
	正常	病理	整体
PA	97.22	95.83	96.53
SFR	100	98.61	99.31
CPP	94.44	86.11	90.28
HNR	93.06	69.44	81.25
Jitter	91.67	62.50	77.08
Shimmer	97.22	63.69	80.56

HNR、Jitter、Shimmer 对正常语音与病理语音同样有一定的判别能力, 但三者的病理分类准确率普遍偏低, 整体分类准确率明显低于 PA、SFR 和 CPP, 说明 HNR、Jitter、Shimmer 的分类识别效果没有 PA 和 SFR、CPP 好, 进一步证明了 PA、SFR 和 CPP 特征参数分类正常与病理语音的优势。

另外, 与原始语音和基音迭代自适应法求得的声门波的分类效果相比较, 结果如表 2 所示, 本文残差信号的 PA 和 SFR 的分类准确率更高, 可以更

准确地诊断出正常语音与病理语音，这与残差信号自身的波形结构特点有关，故残差信号的 PA 和 SFR 在临床评估病理声音方面具有更大的潜在价值；与文献[8]中所用的分类方法相比，本文选择支持向量机分类，预测准确率可以高达 99%，为了增加说服力，同时与学习向量量化(Learning Vector Quantization, LVQ)神经网络相比<sup>[14]</sup>，根据经验知识设置 LVQ 的竞争层神经元个数<sup>[15]</sup>，训练集和测试集也是 2:1 随机生成，运行 10 次后取平均分类率，结果如表 3 所示，说明在本文正常语音与病理语音二分类问题处理中，选用支持向量机的效果相对更优。

表 2 不同信号的 SFR 和 PA 的分类准确率  
Table 2 Classification accuracy of SFR and PA of different signals

特征参数	三种信号的整体分类准确率/%		
	残差信号	原始语音	声门波
SFR	99.31	70.14	65.29
PA	96.53	59.03	64.58

表 3 不同分类方法的分类准确率  
Table 3 Classification accuracy of different classification methods

分类方法	三个特征参数的整体分类准确率/%		
	SFR	PA	CPP
文献[8]的方法	97.93	94.37	85.15
LVQ	96.75	94.43	86.77
SVM	99.31	96.53	90.28

## 5 结论

本文在残差信号的基础上讨论了基音幅值 PA 与频谱平坦度 SFR 和语音倒谱峰值突出 CPP 3 个特征参数。其中，SFR 和 PA 都是基于线性预测模型计算的，二者的测量均不受基频  $F_0$  的影响，而  $F_0$  估计过程中的任何误差都会导致声带噪声测量偏差，所以独立于基频的残差信号的 PA 和 SFR 的分类效果明显优于依赖  $F_0$  测量的特征参数。其中 SFR 能够很好地捕捉声音的噪声特性，几乎可以无重叠地分离正常语音和病理语音，这可以作为语音病理学的有效指标。

相比表征语音信号的其他传统特征参数而言，残差信号的 PA、SFR 和原始语音 CPP 在正常语音与病理语音二分类中效果更好，这充分说明本文选用这 3 个特征参数的有效性和优越性。本文采用 SVM 分类器比其它两个分类器得到的分类准确率更高，说明在分类正常语音与病理语音时，选用支持向量机效果相对更好。本文的研究结果对于临床嗓音学具有重要的指导意义，下一步的工作是病理语音类型的区分，这将对临床实践具有重要价值。

## 参 考 文 献

- [1] 常静雅. 小波域多重分形和能量谱参数的病理嗓音识别研究[D]. 苏州: 苏州大学, 2016.  
CHANG Jingya. Pathological voice recognition study by wavelet domain multifractal and energy spectrum parameters[D]. Suzhou: Soochow University, 2016.
- [2] 张满彩. 病理语音的特征提取与分类研究[D]. 哈尔滨: 哈尔滨工业大学, 2014.  
ZHANG Mancai. Research on feature extraction and classification of pathological speech[D]. Harbin: Harbin Institute of Technology, 2014.
- [3] KOHLER M, MENDOZA L A F, LAZO J G, et al. Classification of voice pathologies using glottal signal parameters[C]//10. Congresso Brasileiro de Inteligência Computacional. 2016: 1-8.
- [4] FORERO A M, KOHLER M, VELLASCO M M, et al. Analysis and classification of voice pathologies using glottal signal parameters[J]. Journal of Voice, 2016, 30(5): 549-556.
- [5] ALKU P. Glottal inverse filtering analysis of human voice production-A review of estimation and parameterization methods of the glottal excitation and their applications[J]. Sadhana, 2011, 36(5): 623-650.
- [6] 张凯晗. 基于逆滤波提取声门波方法的研究与实现[D]. 广州: 暨南大学, 2015.  
ZHANG Kaihan. Research and implementation on the methods of extracting glottal wave based on inverse filtering[D]. Guangzhou: Jinan University, 2015.
- [7] KREIMAN J, GERRATT B R, Antoñanzas-Barroso N. Measures of the glottal source spectrum[J]. Journal of Speech Language & Hearing Research Jslhr, 2007, 50(3): 595-610.
- [8] PARSA V, JAMIESON D G. Identification of pathological voices using glottal noise measures[J]. J Speech Lang Hear Res, 2000, 43(2): 469-485.
- [9] CASTELLANA A, CARULLO A, CORBELLINI S, et al. Discriminating Pathological Voice From Healthy Voice Using Cepstral Peak Prominence Smoothed Distribution in Sustained Vowel[J]. IEEE Transactions on Instrumentation & Measurement, 2018, 67(3): 646-654.
- [10] 李宁. 基于声学参数和支持向量机的病理嗓音分类研究[D]. 上海: 华东师范大学, 2013.  
LI Ning. Automatic Classification for Pathological Voice based on Acoustic Parameters and SVM[D]. Shanghai: East China Normal University, 2013.
- [11] ALNASHERI A, MUHAMMAD G, ALSULAIMAN M, et al. Investigation of voice pathology detection and classification on different frequency regions using correlation functions[J]. Journal of Voice Official Journal of the Voice Foundation, 2016, 31(1): 3.
- [12] SRINIVASAN V, RAMALINGAM V, SELLAM V. Classification of normal and pathological voice using GA and SVM[J]. International Journal of Computer Applications, 2012, 5(1): 1-7.
- [13] 史峰. MATLAB 智能算法 30 个案例分析[M]. 北京: 北京航空航天大学出版社, 2011.  
SHI Feng. Analysis of 30 cases analysis of MATLAB intelligent algorithm[M]. Beijing: Beihang University Press, 2011.
- [14] WEI Y H, LIU W, YANG Y J, et al. A model of gear fault diagnosis based on manifold learning and LVQ[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2018.
- [15] 王小川, 史峰, 郁磊, 等. MATLAB 神经网络 43 个案例分析[M]. 北京: 北京航空航天大学出版社, 2013.  
WANG Xiaochuan, SHI Feng, YU Lei, et al. Analysis of 43 cases analysis of MATLAB neural network [M]. Beijing: Beihang University Press, 2013.