DOI: 10.16300/j.cnki.1000-3630.23061201 CSTR: 32055.14.sxjs.1000-3630.23061201
引用格式：朱婷, 段淑斐, DINGAM Camille, 等 . 结合 EMD 和 FWHT 的构音障碍语音特征增强算法[J]. 声学技术, 2025, **44**(2): 239-251.
[ZHU Ting, DUAN Shufei, DINGAM Camille, et al. Dysarthria speech feature enhancement algorithm by combining empirical mode decomposition and fast Walsh-Hadamard transform[J]. Technical Acoustics, 2025, **44**(2): 239-251.]

结合 EMD 和 FWHT 的构音障碍语音特征增强算法

朱 婷[1]，段淑斐[1]，DINGAM Camille[1]，梁慧芝[2]，张 卫[3]
(1. 太原理工大学电子信息与光学工程学院, 山西太原 030024；2. 英国纽卡斯尔大学计算机学院, 英国纽卡斯尔 NE1 7RU；
3. 北京大学第一医院太原医院, 山西太原 030032)

摘要：传统声学特征易忽略语音的非线性、非平稳特性并且不能同时提取患者声道、声带的病理特性，导致识别模型性能不佳。因此文章提出了一种结合经验模态分解和快速沃尔什-哈达玛变换的构音障碍语音特征增强算法。首先，采用快速傅里叶变换处理语音后，引入经验模态分解自适应提取其本征模态函数；其次，进行快速沃尔什-哈达玛变换；接着，提取基于本征模态函数的统计学特征以及功率谱密度、伽马通频率倒谱系数的增强特征；最后，在 UA Speech 和 TORGO 数据库上进行病情分级研究，并引入了非平衡分类算法评估。结果表明，该算法对比传统特征在病理语音分级研究上是有效的，在考虑类间不平衡后，识别准确率至少提高了 12.18 个百分点。由此，该算法可以更充分表征构音障碍语音特性，对其非平衡性、非线性特性及缺乏同时表征声带和声道中局部病理信息的问题具有一定的改善作用。

关键词：构音障碍；特征增强；经验模态分解；沃尔什-哈达玛变换；病理语音
中图分类号：TN912.3；R767.92 **文献标志码**：A **文章编号**：1000-3630(2025)-02-0239-13

Dysarthria speech feature enhancement algorithm by combining empirical mode decomposition and fast Walsh-Hadamard transform

ZHU Ting[1], DUAN Shufei[1], DINGAM Camille[1], LIANG Huizhi[2], ZHANG Wei[3]
(1. *College of Electronic Information and Optical Engineering*, *Taiyuan University of Technology*, *Taiyuan* 030024, *Shanxi*, *China*; 2. *School of Computing*, *Newcastle University*, *Newcastle NE*1 7*RU*, *UK*; 3. *Taiyuan Hospital*, *Peking University First Hospital*, *Taiyuan* 030032, *Shanxi*, *China*)

Abstract: Dysarthria speech contains the pathological characteristics of the vocal tract and vocal folds. However, these characteristics have not yet been included in traditional acoustic features. Furthermore, the nonlinearity and non-stationarity of speech are also ignored. Therefore, this paper proposes a feature enhancement algorithm for dysarthria speech called WHFEMD by combining empirical mode decomposition (EMD) and fast Walsh-Hadamard transform (FWHT). In this proposed algorithm, the dysarthria speech undergoes fast Fourier transform first, followed by EMD to obtain intrinsic mode functions (IMFs). Then FWHT is applied to generate new coefficients and extract statistical features as well as enhanced features based on Power Spectral Density and Gammatone Frequency Cepstral Coefficients based on IMFs. Disease classification is conducted using data from UA Speech and TORGO databases, which is further evaluated by using an imbalanced classification algorithm. According to experimental findings, WHFEMD enhanced features are significantly superior to traditional features. After balancing the data with the imbalanced classification algorithm, the identification accuracy rate increased by at least 12.18 percentage. This demonstrates that WHFEMD can more ccomprehensively characterize dysarthria speech while addressing issues related to its non-stationary and non-linear characteristics as well as lack of simultaneous characterization of local pathological information in both vocal folds and vocal tracts.

Key words: dysarthria; feature enhancement; empirical mode decomposition; Walsh-Hadamard transform; pathological speech

收稿日期：2023-06-12; **修回日期**：2023-08-29
基金项目：国家自然科学基金青年科学基金 (12004275)、山西省应用基础研究计划面上自然基金 (20210302123186、山西省留学人员科技活动择优资助项目 (20200017)
作者简介：朱婷 (2000—)，女， 重庆人， 硕士生， 研究方向为病理语音， 机器识别。
通信作者：段淑斐， E-mail: duanshufei@tyut.edu.cn

0 引 言

构音障碍 (Dysarthria) 又称“神经性言语障碍”[1]，是由于发音相关的肌肉、器官麻痹或运动

不协调，导致声带、声道、口腔等声学元件功能异常[2]，从而使得说话人的发音韵律准确度严重下降，常见于创伤性脑损伤等疾病中[3]。该疾病反映了构音障碍患者在言语产生过程中呼吸、构音等方面运动的幅度、速度、稳定性等出现异常。语言治疗是常见的治疗手段之一[4]，利用病理语音研究构音障碍语音特点、发病机制等，有助于言语治疗师进行系统评估和康复治疗。其中，提取具有良好病理区分性的声学特征是该识别系统的关键。常见的病理语音声学特征有音高 (pitch)、梅尔频率倒谱系数 (Mel frequency cepstrum coefficients, MFCC)、线性预测系数 (linear prediction coefficients, LPC) 等。一些研究者将语音产生系统视为线性模型，提取不同特征来研究病理嗓音的规律性及其与正常人的差异性，或用来检测构音障碍严重程度[5-7]。以上研究表明了传统声学特征在病理语音检测上的有效性。线性模型将语音视为源滤波器系统的输出。考虑到语音信号的非平稳特性，在特征提取或信号处理上常常会进行分帧、加窗等预处理操作，旨在将语音信号切割成短时平稳的片段。这样的处理方式虽然能在一定程度上捕捉到语音信号的时变特性，但可能会造成信号丢失、帧之间不连续等问题。因此，对于语音信号的分析和处理，研究者需要更加全面地考虑信号的非线性和非平稳特性，以更好地理解和利用其信息。

语音信号的非线性[8]是由信号的产生机制以及言语系统的非线性性质导致的，具体体现在言语传输过程中的线性失真、发声部位的非线性滤波和人耳听觉系统在信号提取过程中的非线性感知。同时，语音作为一种非平稳信号[9]，因说话人口型变换、呼吸调整等导致其时频特性发生了变化。构音障碍患者由于声带、口腔等器官的病变，其语音信号可能存在明显的非线性振动异常。如果忽略这种混沌现象，将会对模型性能产生影响，进而更难以有效区分病理语音与正常语音。研究者们受混沌行为的影响，尝试将语音产生视为一个非线性动态系统来设计非线性模型。Vieira 等[10]提出非平稳性指数作为病理语音的自适应分割方法，相较于线性系统，该方法将病理语音识别的准确率提高了 18 个百分点。Gour 等[11]提取非新型参数使用支持向量机进行病情分级，识别准确率为 73%。因此，选取适当的非平稳、非线性处理技术是从病理语音信号中准确提取特征信息的关键。

在构音障碍语音的研究中发现病理语音发音特性体现在声道、声带等多个声学发音方面且含有相当数量的病理信息。目前常用的声学特征只能反映单方面信息，如 MFCC、LPC、共振峰等主要反映声道信息，基频则反映了声带的运动状态。基于此，寻找能更全面获取构音障碍语音特性的特征参数，以产生更准确有效的性能指标是本研究的关注点。经验模态分解 (empirical mode decomposition, EMD) 作为一种自适应时频信号分析方法[12]，能反映原始信号不同时间尺度的局部特征信息，准确表达信号的瞬时频率特征[13-14]，且具有高的时频聚焦性和信噪比，因此特别适用于非线性、非平稳信号。如 Krishnan 等[15]使用 EMD 从语音信号中提取非线性特征用于识别七种不同的情绪；Rueda 等[16]应用 EMD 的二元滤波器组特性提取发音特征用于识别帕金森语音。Zhang 等[17]提出结合 EMD 和语音信号能谱分析的特征提取方法来区分帕金森病患者和健康人。Hammami 等[18]提出了一种基于经验模态分解和离散小波变换分析来提取病理语音高阶统计特征的新方法。本文将 EMD 应用在构音障碍语音的非线性、非平稳特性分解上，通过自适应分解构建病理信息，从而充分反映声道、声带中的局部特征信息。

快速沃尔什-哈达玛变换 (fast Walsh-Hadamard transform, FWHT) 作为一种求解线性微分方程和积分方程的快速算法[19]，通过分解不同振荡频率下的方波信号，可快速分析语音信号、图像处理中的频谱，是信号变换的理想选择[20]。Prasanna 等[21]将其用于分解脑电信号并提取了一系列非线性特征；Helal 等[22]将沃尔什-哈达玛变换与奇异值分解应用于医学水印图像，得到抗噪声的鲁棒结果；Mohsen 等[23]借助 FWHT 技术分析癫痫脑电图，有助于从非癫痫脑电图中有效推导癫痫发作脑电图。然而，关于 FWHT 的既有研究主要集中于图像处理，目前尚未发现 FWHT 被用于分析研究病理语音信号。该算法的时间复杂度远低于传统的傅里叶变换，这对于处理病理语音这类医学领域数据具有高效性；其次，病理语音信号的频谱通常具有较高能量集中在少数频率分量上的特点，FWHT 作为频域处理方法，有助于分析病理语音信号中的异常频率分量。同时，其信息压缩和并行计算能力可有效提高传输效率和处理速度。所以，本文引入该算法处理病理语音信号的异常波动，进一步加快信号的重建速度，可以自适应地构建构音障碍语音的增强特性。

因此，针对病理语音的非平稳、非线性特性问题以及目前缺少同时表征构音障碍患者声带和声道中局部病理信息的特征参数的局限性，本文将 FWHT 和 EMD 算法结合，自适应地构建了一种构

音障碍语音特征增强算法。该算法主要由快速傅里叶变换 (fast Fourier transform, FFT)、EMD 和 FWHT 三部分组成，进一步增强构音障碍语音识别系统的识别性能。

1　WHFEMD 方法及识别系统

本文提出的病理语音识别方法总体框图如图 1 所示，主要包括原始病理语音信号、基于经验模态分解和快速沃尔什-哈达玛变换的语音处理及分解、特征提取、识别分类等。将本文提出的算法命名为 WHFEMD，其中 WH 指快速沃尔什-哈达玛变换、FEMD 指将语音进行快速傅里叶变换 (FFT) 及经验模态分解 (EMD) 处理的步骤。首先利用 EMD 提取原始语音中的非平稳、非线性特性信息以及声道声带中的局部特征信息，获得其本征模态函数 (intrinsic mode function, IMF)，然后通过快速沃尔什-哈达玛变换获取变换系数，从而提高抗噪性能和运算速度，完成构音障碍语音增强，以达到病理语音特征增强的目的。

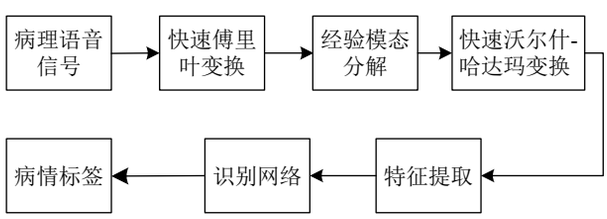


图 1　病理语音识别系统框图

Fig.1　Block diagram of pathology speech recognition system

1.1　经验模态分解

EMD 是通过数据本身的时间尺度特征对给定信号进行逐级自适应分解，从而得到数个本征模态函数及一个剩余分量。对于本征模态函数要求满足：(1) 整个信号序列中极值点的个数h和过零点的个数n必须满足$|h-n|\leqslant 1$；(2) 任意时刻，信号的上下包络线的值v和l必须满足$\frac{1}{2}(v+l)=0$。假设给定一个信号$x(t)$，其 EMD 公式为

$$x(t)=\sum_{i=1}^{n}\mathrm{IMF}_i(t)+r_n(t) \tag{1}$$

式中：$\mathrm{IMF}_i(t)$是信号的本征模态函数，$r_n(t)$为残差。

本征模态函数作为语音信号的基本组成部分，本文通过引入 EMD 算法实现构音障碍语音信号的分解。考虑到病理语音的不规则性，对其局部特征进行自适应分析可以更好表征病理语音的特征；同时，本征模态函数 IMFs 包含所有与说话密切相关的共振峰频率、声带和声道信息。从信号处理角度，高频成分主要反映了信号的瞬时变化和波动，高频 IMF 中标准差可能更敏感；低频成分主要反映了信号的长期变化和趋势，低频 IMF 的方差传达的信息可能更丰富。因此，EMD 所获得频率等语音信息和可以较为全面地代表构音障碍语音的特性，特别是患者所携带的病理信息，从而更好分析病理语音产生机制。

1.2　快速沃尔什-哈达玛变换

对于给定信号$x(t)$及其本征模态函数$\mathrm{IMF}_i(t)$，其 FWHT 变换系数可以定义为

$$\mathrm{FWHT}(u)=\frac{1}{N}\sum_{u=0}^{N-1}\mathrm{IMF}_i(t)\prod_{k=0}^{j-1}(-1)^{b_k(u)b_{(j-1-k)}(u)} \tag{2}$$

式中：$\mathrm{FWHT}(u)$表示沃尔什-哈达玛变换系数；$b_k(u)$表示u的二进制数的第$k+1$位的值；$b_{(j-1-k)}(u)$表示$j-k$位的值；$u=0,1,2,\cdots,N-1$表示当前第u个数据点；N为采样点数，一般为 2 的整数次幂；j为本征模态函数的数量；k为整数。

FWHT 使用±1 离散值代替 FFT 变换中的复指数函数，只进行加减运算，具有计算复杂度低、运算速度快等优点。另外，其压缩能量的方式通常是第一个系数包含了大部分信号能量，这使得第一个系数可很大程度代表原始信号。

图 2(a) 显示了构音障碍语音前 5 阶 IMF 的频谱，从中可以观察到，IMF 的频谱代表了不同语音部分，IMF_1 包含承载了信号的大部分能量，随着 IMF 阶数增加，其频谱近似以 2 的幂次衰减。图 2(b) 显示了其 FWHT 系数，大部分信号能量通常存在于第一序列，也就是 FWHT 压缩能量的方式通常是第一个系数包含大部分的信号能量。

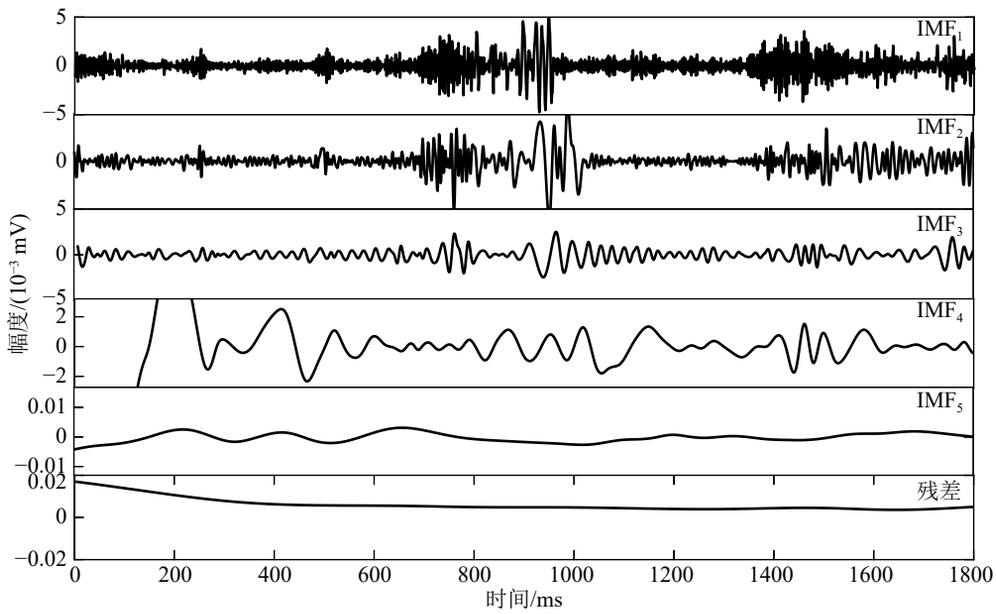
1.3　WHFEMD 算法及特征提取

WHFEMD 算法主要分为信号分解和特征提取。首先利用 FFT 和 EMD 对语音进行处理，得到 FEMD；然后进行 FWHT 变换得到 WHFEMD。分别提取 FEMD、WHFEMD 的统计特征和常见声学特征作为后续分类的数据输入。步骤如下：

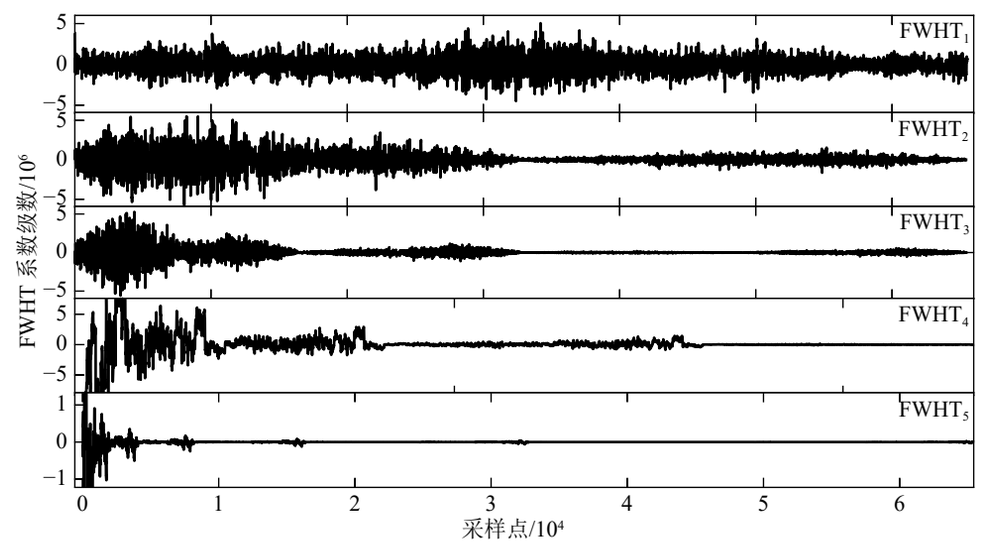
(1) 对原始信号进行 FFT 得到频谱。用$S(f)$和$P(f)$分别表示时间信号的傅里叶变换以及 FFT 后得到的频谱幅度。

(2) 在 FFT 频谱幅度$P(f)$上引入 EMD，得到本征模态函数。需要注意的是，在 EMD 分解过程中，若第j个本征模态函数$\mathrm{IMF}_j(f)$对应的残差$r_j(f)$已经变成单调函数或常数，则不能再进行分解，至此整个分解过程结束。$P(f)$EMD 的具体函数表达式为

$$P(f)=\sum_{i=1}^{j}\mathrm{IMF}_i(f)+r_j(f) \tag{3}$$



(a) 构音障碍语音的 EMD 分解图



(b) 构音障碍语音的沃尔什-哈达玛变换系数

图 2 构音障碍语音的 EMD 分解结果及 FWHT 变换系数

Fig.2 EMD results and FWHT coefficients of dysarthria speech

式中：$\mathrm{IMF}_i(f)$表示第i阶本征模态函数，$i=1,2,3,\cdots,j$；残差$r_j(f)$表示信号的平均趋势，是一个常数或单调序列，j表示本征模态函数的数量。本文经过一系列病情分级实验，选取不同 IMF 阶数时的识别率 (识别率取整数) 如表 1 所示，当 IMF 阶数取 1~5 时，TORGO 数据库上识别准确率最高，为 71%，同时 UA Speech 数据库上的识别率也达到最高，为 44%。因此，本文实验部分选择$j=5$。

(3) 对前 5 阶模态函数运用 FWHT 构建 FEMD 信号，从而得到其变换系数。

(4) 从 FEMD 和 WHFEMD 中提取统计学特征、功率谱密度 (power spectrum density, PSD) 和伽玛通频率倒谱系数 (Gammatone frequency cepstral coefficients, GFCC)。

表 1 IMF 级数的选取 (识别率取整数)

Table 1 Selection of IMF series (integer recognition rate)

IMF阶数	识别率/%	
	TORGO	UA Speech
1~4	65	38
1~5	71	44
1~6	69	44
1~7	69	43
1~8	66	41
1~9	63	40
1~10	62	40

① 统计学特征 (FESF/WHFESF)：考虑到患者与正常人的语音在时域波形上存在较大差异，从 FEMD 和 WHFEMD 中提取每阶 IMF 的统计学特

征，并组合形成维度为$j\times a$的统计学特征向量，即 FESF 和 WHFESF，其中j和a为整数，a表示统计学特征的数量。本文实验取$a=5$，具体包括均值 (mean)、标准差 (standard deviation, SD)、最大值 (maximum)、最小值 (minimum) 和方差 (variance)，从而得到一个 25 维的统计学特征。方差反映信号强弱的变化程度，描述语音的时长不稳定性。标准差关注信号点与平均值的差异，有利于捕捉语音的波动规律。同时研究表明[24-25]，构音障碍患者较健康人构音相关器官和肌肉运动的持续时间延长、幅度减小，速度变慢以及运动轨迹变化更不规则。因此，提取最大值和最小值特征有助于捕捉与异常发音相关的声音信号特性。本文采用了 Z-score 标准化方法来确保不同语料下声音信号在相似尺度上进行分析，以减小声音幅度差异对数据准确性的影响。该归一化方法将特征值转换为相对样本集的均值和标准差的偏差分数，使得不同语料声音幅度以具有相似分布的数据值存在，从而减小了差异性，同时保留了声音幅度的相对变化，因此对于捕捉不同构音障碍患者语音的非线性振动和异常情况仍然有效。表 2 中的随机森林 (random forest, RF) 分类器的病情分级实验结果表明，无论采用 IMF 还是 FWHT 变换后提取标准差和方差，TORGO 和 UA Speech 数据库上识别率至少分别提升了 0.49 和 0.52 个百分点，未出现因方差和标准差数学关联过强引发识别率降低的情况。本文对构音障碍语音提取的方差和标准差能够捕捉语音的不同特性，并提供了区分构音障碍患者的有益信息。

② PSD 增强特征 (FEPSD/WH2FEPSD)：对于语音这样的随机信号，研究其功率谱密度是非常有效的。本文采用周期图法中的 Welch 方法进行谱估计，以提取构音障碍语音的功率谱密度。首先将输入信号分段，再对分段信号进行傅里叶变换以估计功率谱密度，最后将这些估计值平均以获得最终的功率谱估计，其中分段长度决定了频谱的分辨率，重叠率则影响频谱估计的平滑程度。Welch 方法的优势在于可灵活调整窗口大小和重叠程度，允许对信号的局部频率特性进行精细估计，有助于处理非平稳信号。因此，在本实验中首先对各阶 IMF 分段、加窗，得到各分段信号对应的周期图后求取平均。窗函数设置为汉宁窗，离散傅里叶变换点数为 256，重叠率为 50%。本研究中采用 FEMD 来增强 PSD 的性能，得到新特征 FEPSD；另外将 WHFESF 与 FEPSD 结合形成新特征 WH2F-EPSD(流程见图 3)。该特征在对病理语音进行 FFT 变换、EMD 分解得到 IMFs 后，主要分为两部分：1) 对 IMF 进行 FWHT 变换后得到的变换系数求取统计值，形成了特征集合“WHFESF”；

表 2　方差和标准差特征的有效性验证

Table 2　Validation of variance and standard deviation features

数据库	IMF		FWHT	
	特征	识别率/%	特征	识别率/%
TORGO	仅标准差	78.97	仅标准差	80.58
	仅方差	78.90	仅方差	80.06
	FESF	79.45	WHFESF	81.10
UA Speech	仅标准差	57.88	仅标准差	62.45
	仅方差	57.39	仅方差	63.05
	FESF	58.37	WHFESF	63.99

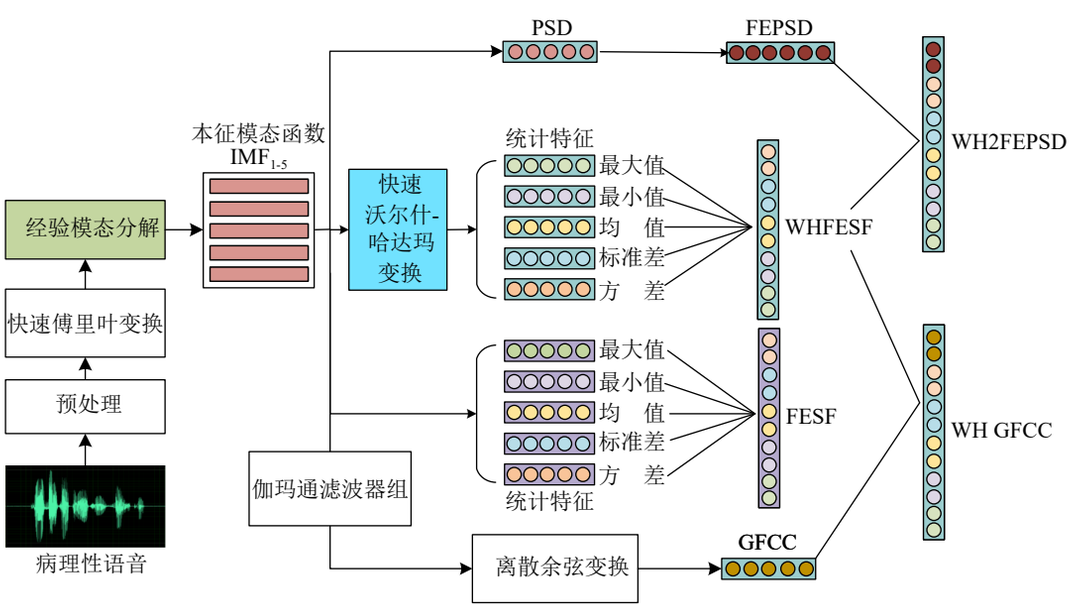


图 3　WHFEMD 算法特征增强流程图

Fig.3　Flow chart of WHFEMD for feature enhancement

2) 分别求取各 IMF 的 PSD 特征值得到特征集合“FEPSD”。最后将这两个特征集合拼接融合从而获得新特征 WH2FEPSD。其中 $j=5$，统计学特征个数取 $a=5$，因此该特征为 30 维。在实验中，研究者可根据实验需求等对其更改，从而实现特征维度扩充等目的。welch 函数的表达式为

$$P_{\text{welch}}=\frac{1}{T}\sum_{t=1}^{T}P_t(f) \tag{4}$$

式中：T 是分段信号的段数；$P_t(f)$ 是第 t 段分段信号的周期图。

③ GFCC 的增强特征 (WHGFCC)：GFCC 是基于 MFCC 的一种修正方法，已被应用于病理语音领域[26]。伽马通滤波器是通过一组等效矩形带宽的带通滤波器来模拟人耳基底膜尖锐的滤波特性。MFCC 作为语音信号中最常用的特征，其精度高、复杂度低，但抗噪声能力不高。研究表明[27]，GFCC 特征具有较强的抗噪性能和声学鲁棒性，能有效捕捉语音的本质特征。GFCC 提取方法与 MFCC 相似，主要区别如下：1) 频率尺度，GFCC 是基于等效矩形带宽尺度，而 MFCC 是基于梅尔 (Mel) 尺度，这使得 GFCC 在低频段比 MFCC 提供了更高的分辨率[28]；2) 压缩方法，相比 MFCC 的对数压缩方法，GFCC 的指数压缩方法可更好地模拟听觉系统中的非线性特性，从而具有良好的抗干扰能力。基于此，本文将 EMD 分解和 FWHT 变换后的信号提取的 GFCC 特征与 WHFESF 特征结合，得到新特征参数 WHGFCC 来提高系统的整体测试精度。

④ 常用声学特征：用于对照实验，包括 MFCC、LPC、GFCC、PSD、Pitch、线谱对 (line spectral pairs, LSP)。

2 实验与结果分析

2.1 数据库介绍

UA Speech 数据集[29]：由伊利诺伊大学创建。受试者为 15 名脑瘫患者和 13 名正常人。语料包括数字、字母、计算机命令、常用和不常用的单词等。详情见网址：http://www.isle.illinois.edu/sst/data/UASpeech/。本文共选用 21 188 个来自患者和正常人的单词进行构音障碍病情分级研究，具体数据如表 3 所示。

TORGO 数据集[30]：由多伦多大学创建，收录了 8 名运动型构音障碍患者 (3 名女性，5 名男性) 和 7 名正常人 (3 名女性，4 名男性) 的声学数据、运动学数据。详细信息见官方网址介绍：http://www.cs.toronto.edu/data/TORGO/torgo.html。本文选用短词和限制句录制的声学数据对构音障碍病情分级研究。UA Speech 和 TORGO 数据库数据选择如表 3 所示。

表 3 UA Speech 和 TORGO 数据库数据选择
Table 3 Data selection on UA Speech and TORGO databases

数据库	构音障碍严重程度	样本数	合计
UA Speech	重度	2 861	21 188
	中度	2 280	
	轻度	2 280	
	轻微度	3 825	
	对照组	9 942	
TORGO	重度	194	4 867
	中度	259	
	轻度	596	
	对照组	3 818	

2.2 实验设置

本文采用集成算法中的三种分类器完成下游分类任务，分别是 RF、袋装算法 (bagging, BG) 和多层感知器 (multilayer perceptron, MLP)。RF[29]同时使用袋装和特征选择方法，可消除决策树的限制，避免过拟合来提高精度。BG 算法[31]主要通过从现有的训练集中随机选择替换的数据来构建一个新的重新采样的训练集。本文使用的 BG 采用了决策树后剪枝算法。MLP[32]是一种监督学习算法，由输入层、隐藏层、输出层组成，其层与层之间是完全连接的，可用于分类问题。首先以 7∶3 划分训练集和测试集，然后将常见声学特征 (MFCC、LPC、PSD、Pitch、LSP) 与本文所提算法的特征进行比较，以验证该算法的有效性和可靠性。本文所有实验均使用 HP ZBook Power G7 移动工作站运行 Matlab 2022a 程序，处理器配置为 Intel Core i7-11850H，内存为 32 GiB。

2.3 实验结果

本节首先比较使用 WHFEMD 和 FEMD 提取的统计学特征与常用声学特征的识别结果，然后考虑样本间的不平衡问题，引入非平衡分类算法[31]在本文的特征增强算法上加以应用，进一步评估算法的适应性和鲁棒性。该非平衡分类方法充分考虑类之间的不平衡问题，结合主成分分析、人工少数类过采样法和期望最大化算法对病理语音的严重程度进行分类。并且在 TORGO、UA Speech 数据库上取得了显著优于现有非平衡分类算法的评价结果。

2.3.1 常用声学特征比较

在 UA Speech 数据库上，本文所提方法得到

了差异较为显著的结果如表 4 所示。由表 4 可知，WHFEMD 算法提取的统计学特征 WHFESF 在 RF、BG、MLP 分类器识别准确率分别为 63.99%、62.65%、62.87%，与 PSD、Pitch、LPC、LSP 的结果相比，在 RF 分类器下提高了 14.4 个百分点、13.25 个百分点、8.07 个百分点、11.19 个百分点，在 BG 分类器下分别提高了 12.16 个百分点、8.62 个百分点、7.38 个百分点、8.64 个百分点，在 MLP 分类器下分别得到了 11.72 个百分点、9.05 个百分点、7.6 个百分点、8.86 个百分点。

表 4　FEMD/WHFEMD 所提统计学特征与常用声学特征比较 (UA Speech 数据库)

Table 4　Comparison of statistical features presented by FEMD/WHFEMD and common acoustic features (UA Speech database)

特征 (分类器)	准确率/%	特征 (分类器)	准确率/%	特征 (分类器)	准确率/%
PSD(RF)	49.59	PSD(BG)	50.49	PSD(MLP)	51.15
Pitch(RF)	50.74	Pitch(BG)	54.03	Pitch(MLP)	53.82
LPC(RF)	55.92	LPC(BG)	55.27	LPC(MLP)	55.27
LSP(RF)	52.80	LSP(BG)	54.01	LSP(MLP)	54.01
FESF(RF)	58.37	FESF(BG)	56.53	FESF(MLP)	56.53
WHFESF (RF)	63.99	WHFESF (BG)	62.65	WHFESF (MLP)	62.87

在 TORGO 数据库上，本文方法同样得到了更好的结果。如表 5 所示，WHFEMD 算法所提的统计学特征 WHFESF 识别结果在 RF、BG、MLP 分类器上的准确率分别为 81.10%、80.00% 和 78.84%，明显优于所有其他特征，包括 MFCC 特征。WHFESF 与 PSD、Pitch、LPC、LSP、MFCC 相比，RF 分类准确率分别增加了 3.84 个百分点、2.74 个百分点、2.06 个百分点、2.81 个百分点和 0.35 个百分点，BG 分类准确率分别提高了 2.26 个百分点、0.07 个百分点、2.04 个百分点、1.85 个百分点和 0.68 个百分点，MLP 分类准确率分别提高了 0.55 个百分点、0.07 个百分点、1.1 个百分点、0.69 个百分点和 0.55 个百分点。

表 5　FEMD/WHFEMD 所提统计学特征与常用声学特征比较 (TORGO 数据库)

Table 5　Comparison of statistical features presented by FEMD/WHFEMD and common acoustic features (TORGO database)

特征 (分类器)	准确率/%	特征 (分类器)	准确率/%	特征 (分类器)	准确率/%
PSD(RF)	77.26	PSD(BG)	77.74	PSD(MLP)	78.29
Pitch(RF)	78.36	Pitch(BG)	79.93	Pitch (MLP)	78.77
LPC(RF)	79.04	LPC(BG)	77.96	LPC(MLP)	77.74
LSP(RF)	78.29	LSP(BG)	78.15	LSP(MLP)	78.15
MFCC(RF)	80.75	MFCC(RF)	79.32	MFCC(MLP)	78.29
FESF(RF)	79.45	FESF(BG)	79.38	FESF(MLP)	78.28
WHFESF (RF)	81.10	WHFESF (BG)	80.00	WHFESF (MLP)	78.84

无论是 TORGO 还是 UA Speech 数据库，FESF 和 WHFESF 识别结果总体均优于其他目前研究中普遍使用常用的声学特征。这表明，本文特征增强方法显著提高了分类精度，能正确识别不同程度的构音障碍语音，鲁棒性高，可有效提高构音障碍语音识别系统的性能。

如表 6 所示，对 UA Speech 数据库的数据采用 FEMD 算法提取 PSD 特征所得到的 FEPSD 特征，RF 识别精度从 49.59% 提高到 60.45%，BG 识别精度从 50.49% 提高到 59.33%，MLP 识别精度从 51.15% 提升至 58.84%。应用 WHFEMD 后 (即 WH2FEPSD)，相比 PSD，RF、BG、MLP 分类器识别精度分别提高了 18.14 个百分点、16.16 个百分点、15.5 个百分点。这些改进表明了 WHFEMD 和 FEMD 方法在该特征上的重要性及可行性。

从表 6 中还可以看出，当将 WHFESF 统计学特征与 GFCC 特征组合使用时 (即 WHGFCC)，在 RF 分类器准确率从 76.54% 提升至 80.08%，使用 BG 分类器的识别准确率从 73.25% 提高到了 76.18%，使用 MLP 分类器的识别准确率从 80.02% 提高至 82.50%，这同样表明了 WHFEMD 方法的适用性。并且，以上结果优于包括 MFCC 特征在内的所有常用声学特征分类精度，同时在这一数据集上相比于其他方法具有更优的结果[33]。

表 6　FEMD 和 WHFEMD 应用于 PSD/GFCC 的识别准确率 (UA Speech 数据库)

Table 6　Recognition accuracies of FEMD and WHFEMD applied to PSD/GFCC (UA Speech database)

特征 (分类器)	准确率/%	特征 (分类器)	准确率/%	特征 (分类器)	准确率/%
PSD(RF)	49.59	PSD(BG)	50.49	PSD(MLP)	51.15
MFCC(RF)	67.10	MFCC(BG)	65.10	MFCC(MLP)	65.10
FEPSD(RF)	60.45	FEPSD(BG)	59.33	FEPSD(MLP)	58.84
WH2FEPSD (RF)	67.73	WH2FEPSD (BG)	66.65	WH2FEPSD (MLP)	66.65
GFCC(RF)	76.54	GFCC(BG)	73.25	GFCC(MLP)	80.02
WHGFCC (RF)	80.08	WHGFCC (BG)	76.18	WHGFCC (MLP)	82.50

如表 7 所示在 TORGO 数据库上，将 FEMD 和 WHFEMD 应用于 PSD 和 GFCC 上同样提高了其识别的性能。PSD 特征上的识别准确率分别提升了 4.38 个百分点、3.15 个百分点、2.6 个百分点, GFCC 特征的分类准确率也相对增加了 1.57 个百分点、2.67 个百分点和 3.8 个百分点。对比之前在

表 7 FEMD 和 WHFEMD 应用于 PSD/GFCC 的识别准确率 (TORGO 数据库)

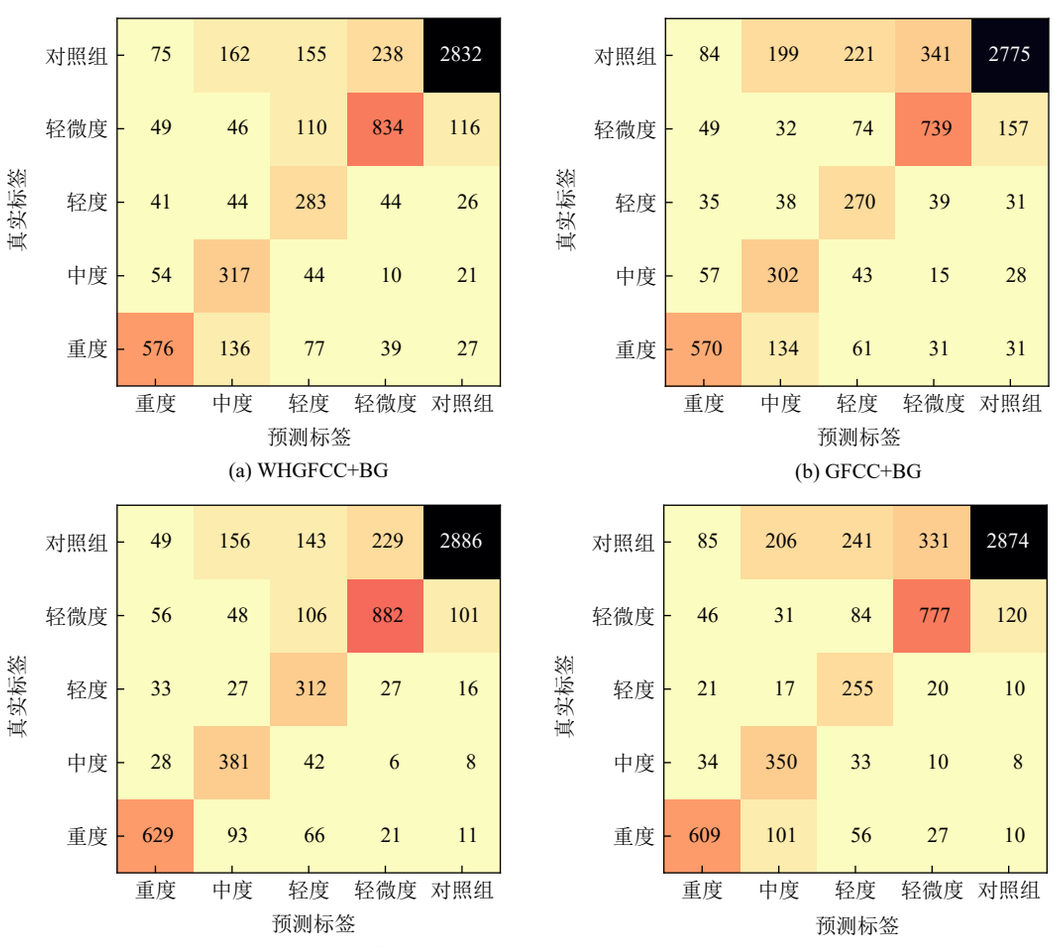
Table 7 Recognition accuracies of FEMD and WHFEMD applied to PSD/GFCC (TORGO database)

特征 (分类器)	准确率/%	特征 (分类器)	准确率/%	特征 (分类器)	准确率/%
PSD(RF)	77.26	PSD(BG)	77.74	PSD(MLP)	78.29
MFCC(RF)	80.75	MFCC(BG)	79.32	MFCC(MLP)	78.28
FEPSD(RF)	79.52	FEPSD(BG)	79.38	FEPSD(MLP)	79.38
WH2FEPSD (RF)	81.64	WH2FEPSD (BG)	80.89	WH2FEPSD (MLP)	80.89
GFCC(RF)	85.14	GFCC(BG)	83.63	GFCC(MLP)	86.78
WHGFCC (RF)	86.71	WHGFCC (BG)	86.3	WHGFCC (MLP)	90.58

TORGO 数据库上使用相同语料类型进行的分类识别工作，本文所提方法得到了更精确的结果[25]。总之，与其他常用特征相比，本文所提的特征增强算法在识别不同程度的构音障碍语音和正常语音方面具有良好的可行性和有效性。

下面将从混淆矩阵的角度，进一步分析 GFCC 和 PSD 在两个数据库上的表现。以 RF 和 BG 分类器为例，分别比较使用本文所提出的特征增强算法应用在 PSD 和 GFCC 上得到的 WH2FEPSD 与 WHGFCC 以及传统特征提取算法得到的 PSD 和 GFCC 在病情分级上的性能，结果如图 4 和图 5 所示。

如图 4(a) 中“576”表示使用 WHFEMD 算法后提取的 GFCC 增强特征 WHGFCC 在 UA Speech 数据库上使用 BG 分类器预测“重度”的正确样本数为 576。从图 4 和图 5 可知，基于 UA Speech 数据库的病理语音分类实验中，本文方法在所有类别中均占主导地位。



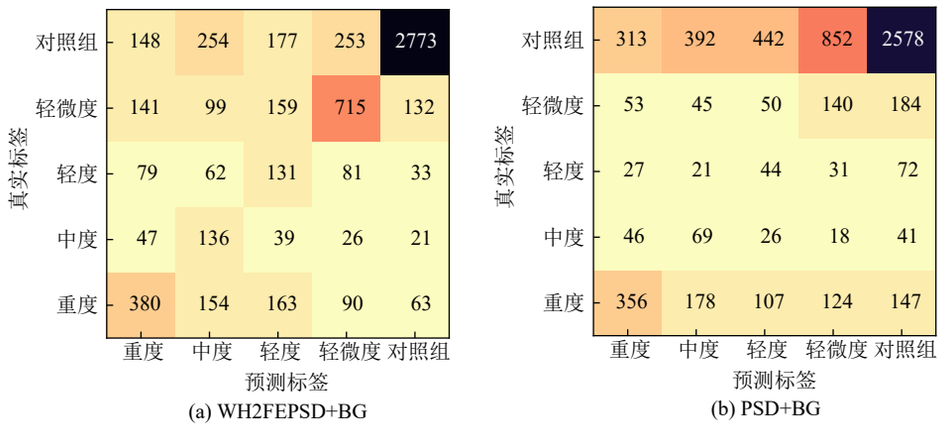
(a) WHGFCC+BG (b) GFCC+BG

(c) WHGFCC+RF (d) GFCC+RF

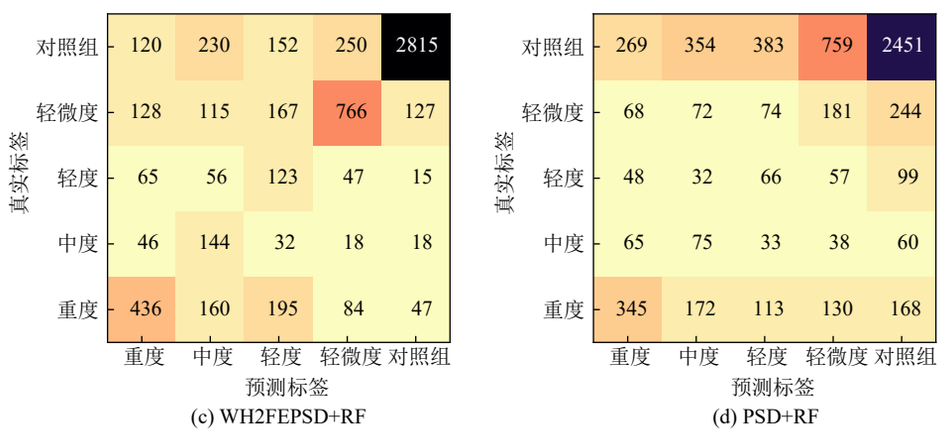
图 4 WHGFCC 和 GFCC 特征在 UA Speech 数据库的混淆矩阵

Fig.4 Confusion matrix of WHGFCC and GFCC features on UA Speech database

当使用 TORGO 数据库时，不同特征在病情分级上的性能情况如图 6~7 所示，尽管在 GFCC 特征上，“重度”构音障碍语音使用本文算法对 GFCC 特征增强后 BG 分类器减少了 3 个正确分类数据；但是与原始特征结果相比，本文所提出的特征增强方法在很大程度上保持了其对构音障碍不同严重程度分类 (即严重、中度和轻度) 的鲁棒性，能够有效减少分类的混淆度。这意味着本文所提方



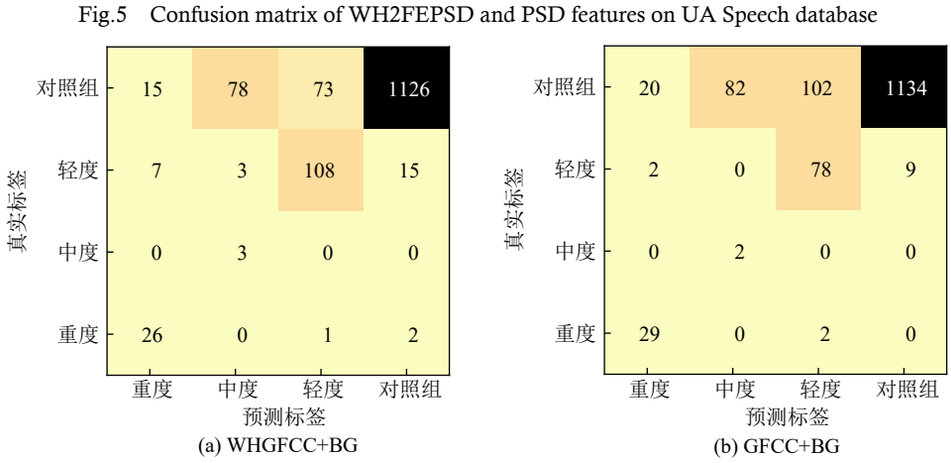
(a) WH2FEPSD+BG　　(b) PSD+BG



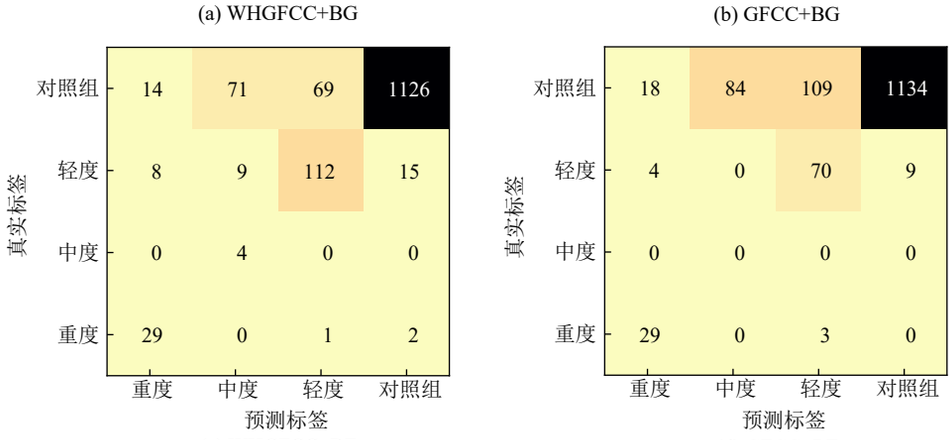
(c) WH2FEPSD+RF　　(d) PSD+RF

图 5　WH2FEPSD 和 PSD 特征在 UA Speech 数据库的混淆矩阵

Fig.5　Confusion matrix of WH2FEPSD and PSD features on UA Speech database



(a) WHGFCC+BG　　(b) GFCC+BG



(c) WHGFCC+RF　　(d) GFCC+RF

图 6　WHGFCC 和 GFCC 特征在 TORGO 数据库的混淆矩阵

Fig.6　Confusion matrix of WHGFCC and GFCC features on TORGO database

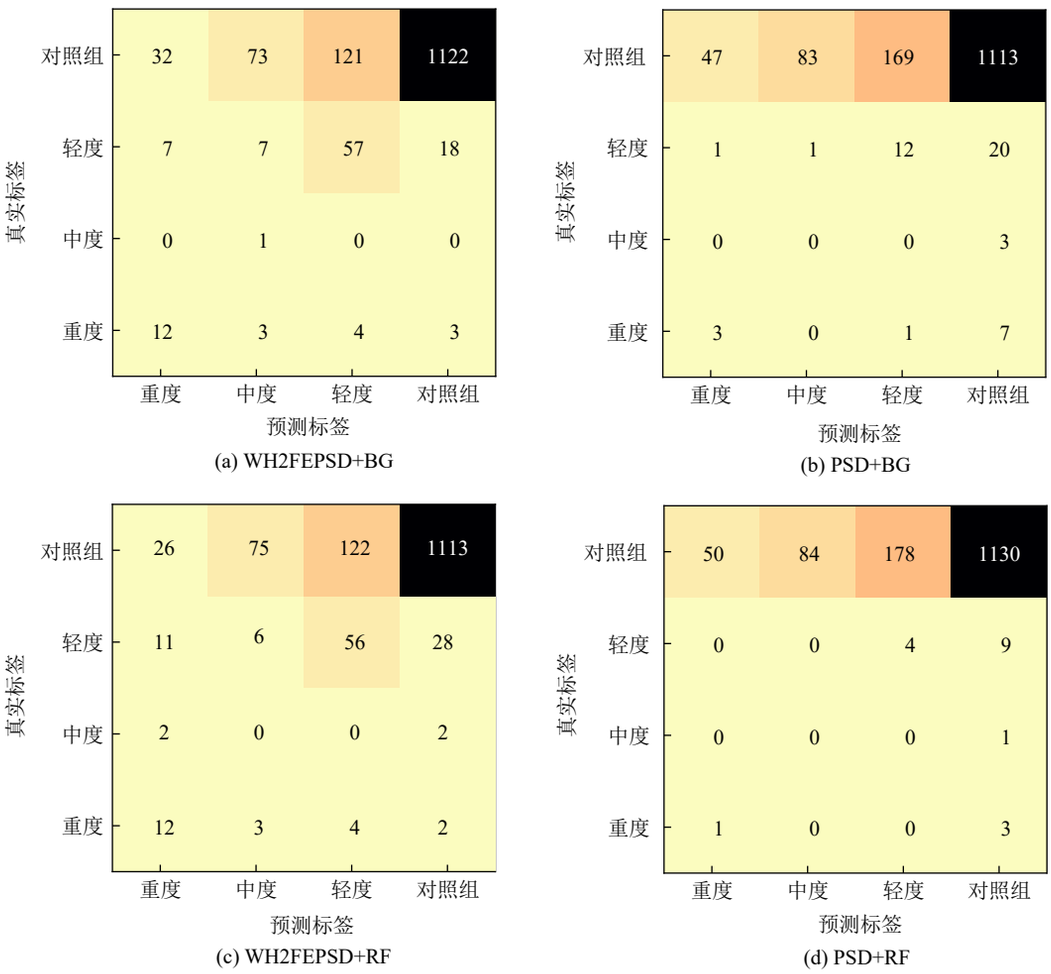


图 7 WH2FEPSD 和 PSD 特征在 TORGO 数据库的混淆矩阵

Fig.7 Confusion matrix of WH2FEPSD and PSD features on TORGO database

法非常适合构音障碍语音，能充分提取出病理语音中携带的病理特性。

2.3.2 WHFEMD 算法和传统算法复杂度比较

传统语音增强算法包括自适应滤波法、谱减法以及基于小波变换相关的语音增强算法等。自适应滤波器法[34]是通过估计噪声的统计特性和语音信号的相关性来调整滤波器的参数，从而实现对噪声的抑制。谱减法[35]则是通过计算语音信号的频谱，并将其噪声减去来实现对噪声的抑制。基于小波包的语音增强算法[36]是一种利用小波包分析的方法。它通过将语音信号分解为不同尺度和频带的小波包系数，并对噪声成分进行抑制和去除，从而减弱噪声。本节比较了本文所提算法 (WHFEMD) 与这三种传统算法的运行时间和复杂度，如表 8 所示。由表 8 可见，本文所提算法在参数量、模型所占空间内存方面与传统算法属于同一数量级，在运行时长上有所增加，时间复杂度有所增加。这主要是由 EMD 分解算法的迭代分解过程引起的，但是根据实验部分的识别准确率结果来看，本文所提算法在增加一部分复杂度的情况下，特征的差异性、区分度有明显提升，总体识别结果更优。因此，WHFEMD 对于构音障碍等病理语音特征增强是可取的。

同时，不同实验的计算复杂度有助于更好地理解和比较不同算法的性能和效率。本文重点在提出构音障碍相关的病理语音的特征增强算法，在实验部分后端分类器相同，因此主要对算法提取不同特征的复杂度和运行时间进行对比。通过参数量可一

表 8 WHFEMD 与传统算法的复杂度和运行时间对比

Table 8 Comparison of complexity and running time BetweenWHFEMD and traditional algorithms

算法	参数量/10^5	模型大小/MiB	运行时长/s	复杂度
自适应滤波法	1.02	4.38	482.56	$O(N)$
谱减法	1.25	4.58	332.72	$O(N+N\log_2 N)$
基于小波的语音增强算法	0.24	9.38	956.89	$O(N+N\log_2 N)$
本文所提算法 (WHFEMD)	1.25	5.58	1 048.49	$O(N^2+N\log_2 N)$

注：N为信号长度。

定程度表示模型的规模，运行时长主要是包括两个数据集所有语音数据特征提取及储存的时长。

如表 9 所示，本文所提算法的统计学特征 FESF 和 WHFESF 参数量的数量级在 10^5，但是运行时间减少了近 305 s。结合表 3~4 的识别准确率结果，本文所提算法在准确率提升的情况下、运行时间减少，这说明该算法对于病理语音识别是高效、准确的。本文算法的 PSD 特征增强 (WH2FEPSD) 作为 WHFESF 和 FEPSD 的融合特征，对比两者运行时长总和减少了 560 s，但是参数量与 FEPSD 基本持恒。同时在 TORGO 和 UA Speech 数据库上的识别准确率 (见表 6~7) 分别提高了至少 2.6 个百分点和 7.69 个百分点。WHGFCC 特征是将 GFCC 和 WHFESF 特征相结合，两个特征运行时长总和相比本次算法运行时长缩短了 502 s，同时识别准确率至少提升了 1.57 个百分点。总体而言，本文所提算法在特征提取的参数量、内存大小上有所增加，但是运行时长和算法识别准确率上均有着显著的提升，不仅保证了特征显著性和识别效果，同时也降低了功耗。

表 9　WHFEMD 提取不同特征的复杂度和运行时间对比
Table 9　Comparison of complexity and running time of different feature extraction with WHFEMD

特征	参数量/10^5	运行时长/s
FESF	1.02	646.32
WHFESF	1.25	951.15
PSD	0.24	580.76
FEPSD	1.25	878.28
WH2FEPSD	1.26	1 269.15
GFCC	0.25	619.78
WHGFCC	1.25	1 068.29

2.3.3　WHFEMD 算法在非平衡分类算法上的应用

由表 4~7 可知，WHFEMD 应用于 GFCC 特征可以得到最优效果，因此本节主要讨论以 WHGFCC 特征为输入，应用非平衡分类方法[31]进行分类的效果，从而进一步检验本文所提的特征增强算法以及之前提出的非平衡分类算法的有效性。

在表 10 中，与原始的支持向量机 (support vector machine，SVM) 和朴素贝叶斯 (naive Bayes, NB) 相比，在 TORGO 数据库上的识别准确率分别提高了 12.18 个百分点和 45.15 个百分点，在 UA Speech 数据库上分别提高了 34.55 个百分点和 15.56 个百分点。这表明，基于支持向量机和朴素贝叶斯的非平衡分类方法明显优于原来的分类器结果。总体上，参考文献[31]所提出的非平衡方法相对于单一的 SVM 和 NB 识别率得到很大的提高，在 TORGO 和 UA Speech 数据集上也得到了验证，并且在将本文的特征增强算法 WHFEMD 应用于 GFCC 特征上得到了比使用的 RF, BG 和 MLP 更为良好的识别性能。以上结果较好地评估了本文所提方法对构音障碍语音分类的美好前景。

表 10　TORGO 和 UA Speech 数据库的识别准确率比较结果
Table 10　Comparison of identification accuracy rates on the TORGO and UA Speech databases

算法	识别准确率/%	
	TORGO 数据库	UA Speech 数据库
SVM[37]	77.84	51.66
非平衡分类器(SVM)[31]	90.02	86.21
NB[37]	40.12	60.30
非平衡分类器(NB)[31]	85.27	75.86

3　结 论

构音障碍作为一种言语障碍，患者会发音迟钝、含糊不清，标准的语音识别系统不能满足构音障碍语音识别需求。针对病理语音的非线性和非平稳特性，以及现有特征参数无法同时包含声道、声带病理信息的问题数，本文提出了一种结合 EMD 和 FWHT 的构音障碍语音特征增强的新方法——WHFEMD，旨在最大程度上地表征病理语音特性，从而提升模型性能。该方法主要由快速傅里叶变换、经验模态分解、快速沃尔什-哈达玛变换三部分组成。针对 UA Speech 和 TORGO 数据库中的语音数据，该组特征参数较为全面地探究了不同严重程度的构音障碍患者以及正常人之间的区分性，证明了引入经验模态分解和快速沃尔什-哈达玛变换实现自适应构建构音障碍语音特征增强算法的可行性，提高了构音障碍语音病情分级研究的科学性和准确性，为之后在不同病理病因方面的应用提供了有力的支撑。本研究的主要创新点如下：(1) 引入 EMD 自适应分解病理语音信号，从而同时捕获并提供了声道和声带的病理局部特征信息；(2) 引入 FWHT 变换方法处理病理语音的异常波动，得以稳健的抗噪性能和信号重建速度；(3) 提取了一组增强的统计学、声学病理语音特征参数。

实验结果表明，WHFEMD 和 FEMD 算法提取的统计学特征对于构音障碍语音可达到较高的识别精度，有效提高了 PSD 和 GFCC 的识别性能。通过 EMD 分解提取病理语音局部特性的关键特征信息，可更深入分析研究病理语音的发声机制，将为临床诊疗提供辅助指导。FWHT 的引入能提取声道声带等病理相关的病理情况，提高运行效率，适

用于实际医学的大规模数据量。未来将结合声学、声门、面部微表情等多模态数据研究构音障碍严重程度评估模型的可行性，进一步从人体发音机理探讨构音障碍病因。

参 考 文 献

[1] 席艳玲, 黄昭鸣. 言语障碍康复治疗技术[M]. 北京: 人民卫生出版社, 2020.

[2] KODRASI I, BOURLARD H. Spectro-temporal sparsity characterization for dysarthric speech detection[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing, 28: 1210-1222.

[3] 庞宇峰, 黄娟, 徐蓓峥, 等. 病态嗓音的定量分析及人工神经网络识别[J]. 临床耳鼻咽喉头颈外科杂志, 2017, **31**(2): 100-102.
PANG Yufeng, HUANG Juan, XU Beizheng, et al. Quantitative analysis of pathological voice and identification with artificial neural network[J]. Journal of Clinical Otorhinolaryngology Head and Neck Surgery, 2017, **31**(2): 100-102.

[4] WANG J L, XU H Y, PENG X Y, et al. Pathological voice classification based on multi-domain features and deep hierarchical extreme learning machine[J]. The Journal of the Acoustical Society of America, 2023, **153**(1): 423.

[5] CHANDRASHEKAR H M, KARJIGI V, SREEDEVI N. Investigation of different time-frequency representations for intelligibility assessment of dysarthric speech[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020, **28**(12): 2880-2889.

[6] WANG S S, WANG C T, LAI C C, et al. Continuous speech for improved learning pathological voice disorders[J]. IEEE Open Journal of Engineering in Medicine and Biology, 2022, **3**: 25-33.

[7] JOSHY A A, RAJAN R. Automated dysarthria severity classification: a study on acoustic features and deep learning techniques[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022, **30**: 1147-1157.

[8] GAO X H. A nonlinear prediction model for Chinese speech signal based on RBF neural network[J]. Multimedia Tools and Applications, 2022, **81**(4): 5033-5049.

[9] KERKENI L, SERRESTOU Y, RAOOF K, et al. Automatic speech emotion recognition using an optimal combination of features based on EMD-TKEO[J]. Speech Communication, 2019, **114**(C): 22-35.

[10] VIEIRA V J D, COSTA S C, CORREIA S E N. Non-stationarity-based adaptive segmentation applied to voice disorder discrimination[J]. IEEE Access, 2023, **11**: 54750-54759.

[11] GOUR G B, UDAYASHANKARA V, BADAKH D K, et al. Quest for speech enhancement method in the analysis of pathological voices[J]. Circuits, Systems, and Signal Processing, 2023, **42**(6): 3617-3648.

[12] GYAMERAH S A. On forecasting the intraday Bitcoin price using ensemble of variational mode decomposition and generalized additive model[J]. Journal of King Saud University - Computer and Information Sciences, 2022, **34**(3): 1003-1009.

[13] KARAN B, SAHU S S, MAHTO K. Parkinson disease prediction using intrinsic mode function based features from speech signal[J]. Biocybernetics and Biomedical Engineering, 2020, **40**(1): 249-264.

[14] AZIZ S, HASSAN NAQVI S Z, KHAN M U, et al. Electricity Theft Detection using Empirical Mode Decomposition and K-Nearest Neighbors[C]//2020 International Conference on Emerging Trends in Smart Technologies (ICETST). Karachi, Pakistan. IEEE, 2020: 1-5.

[15] KRISHNAN P T, JOSEPH RAJ A N, RAJANGAM V. Emotion classification from speech signal based on empirical mode decomposition and non-linear features[J]. Complex & Intelligent Systems, 2021, **7**(4): 1919-1934.

[16] RUEDA A, VÁSQUEZ-CORREA J C, OROZCO-ARROYAVE J R, et al. Empirical mode decomposition articulation feature extraction on Parkinson's Diadochokinesia[J]. Computer Speech & Language, 2022, **72**: 101322.

[17] ZHANG T, ZHANG Y J, SUN H, et al. Parkinson disease detection using energy direction features based on EMD from voice signal[J]. Biocybernetics and Biomedical Engineering, 2021, **41**(1): 127-141.

[18] HAMMAMI I, SALHI L, LABIDI S. Voice pathologies classification and detection using EMD-DWT analysis based on higher order statistic features[J]. IRBM, 2020, **41**(3): 161-171.

[19] YU X, YANG F, GAO B, et al. Deep compressive single pixel imaging by reordering hadamard basis: a comparative study[J]. IEEE Access, 2020, **8**: 55773-55784.

[20] LIU T, LAI S N, YUAN X C, et al. A novel blockchain-watermarking mechanism utilizing interplanetary file system and fast Walsh hadamard transform[J]. iScience, 2024, **27**(9): 110821.

[21] PRASANNA J, SUBATHRA M S P, ABED MOHAMMED M, et al. Detection of focal and non-focal electroencephalogram signals using fast walsh-hadamard transform and artificial neural network[J]. Sensors, 2020, **20**(17): 4952.

[22] HELAL S, SALEM N. A hybrid watermarking scheme using Walsh hadamard transform and SVD[J]. Procedia Computer Science, 2021, **194**(C): 246-254.

[23] MOHSEN S, GHONEIM S S M, ALZAIDI M S, et al. Classification of electroencephalogram signals using LSTM and SVM based on fast walsh-hadamard transform[J]. Computers, Materials & Continua, 2023, **75**(3): 5271-5286.

[24] 宋景, 王勇丽, 万勤, 等. 电磁发音仪在构音障碍患者言语运动康复中的应用进展[J]. 听力学及言语疾病杂志, 2022, **30**(2): 214-218.
SONG Jing, WANG Yongli, WAN Qin, et al. Application progress of electromagnetic phonograph in speech motor rehabilitation of patients with dysarthria[J]. Journal of Audiology and Speech Pathology, 2022, **30**(2): 214-218.

[25] DUAN S F, ZHANG X Y, YAN M M, et al. Statistical distribution exploration of tongue movement for pathological articulation on word/sentence level[J]. IEEE Access, 2020, **8**: 91057-91069.

[26] LAURAITIS A, MASKELIUNAS R, DAMASEVICIUS R, et al. Detection of speech impairments using cepstrum, auditory spectrogram and wavelet time scattering domain features[J]. IEEE Access, 2020, **8**: 96162-96172.

[27] KROBBA A, DEBYECHE M, SELOUANI S A. A novel hybrid feature method based on Caelen auditory model and gammatone filterbank for robust speaker recognition under noisy environment and speech coding distortion[J]. Multimedia Tools and Applications, 2023, **82**(11): 16195-16212.

[28] ISLAM R, ABDEL-RAHEEM E, TARIQUE M. A novel pathological voice identification technique through simulated cochlear implant processing systems[J]. Applied Sciences,

2022, **12**(5): 2398.
[29] TONG H. Automatic assessment of dysarthric severity level using audio-video cross-modal approach in deep learning[D]. Auckland, New Zealand: Auckland University of Technology, 2020.
[30] NARENDRA N P, ALKU P. Glottal source information for pathological voice detection[J]. IEEE Access, 2020, **8**: 67745-67755.
[31] DINGAM C, ZHANG X Y, DUAN S F, et al. Imbalanced data classification of pathological speech using PCA, SMOTE, and expectation maximization[C]//LIANG Q, WANG W, LIU X, et al. International Conference in Communications, Signal Processing, and Systems. Singapore: Springer, 2022: 309-317.
[32] FAKHRZAD F, JOWKAR A, HOSSEINZADEH J. Mathematical modeling and optimizing the *in vitro* shoot proliferation of wallflower using multilayer perceptron non-dominated sorting genetic algorithm-II (MLP-NSGAII)[J]. PLoS One, 2022, **17**(9): e0273009.
[33] MARTÍNEZ D, LLEIDA E, GREEN P, et al. Intelligibility assessment and speech recognizer word accuracy rate prediction for dysarthric speakers in a factor analysis subspace[J]. ACM Transactions on Accessible Computing, **6**(3): 10.
[34] PAULINE S H, DHANALAKSHMI S, KUMAR R, et al. Noise reduction in speech signal of Parkinson's Disease (PD) patients using optimal variable stage cascaded adaptive filter configuration[J]. Biomedical Signal Processing and Control, 2022, **77**: 103802.
[35] CHINCHU M S, KIRUBAGARI B, MATHEW K. Classification of pathological disorders using optimization enabled deep neuro fuzzy network[J]. Biomedical Signal Processing and Control, 2022, **78**: 103771.
[36] JADDA A, PRABHA I S. Speech enhancement via adaptive Wiener filtering and optimized deep learning framework[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2023, **21**(1): 2250032.
[37] GU J, LU S. An effective intrusion detection approach using SVM with naïve Bayes feature embedding[J]. Computers & Security, 2021, **103**: 102158.

•简 讯•

《声学技术》编委会换届的决定

《声学技术》第五届编委会已经到期，中国科学院声学研究所东海研究站、同济大学、上海市声学学会、中国船舶集团有限公司第七二六研究所四个主办单位领导于 2025 年 4 月商讨并征得了有关专家的同意，对第六届编委会的组成方案做出以下几项决定：

(1) 第六届编委会任期 5 年 (2025-2030 年)；

(2) 聘请中国科学院声学研究所东海研究站胡长青任主编；

(3) 聘请杨德森、成利、杨益新、项延训、毛东兴、杜选民、郑元义任副主编，聘请李秀红任常务副主编；

(4) 编委会委员根据学科分布和单位分布，由主编与主办单位领导商讨聘请。

《声学技术》编辑部