

# 腭裂康复手术客观评价的研究 ——应用塞音检测的方法\*

胡 琦<sup>1,3</sup> 赵庆卫<sup>1†</sup> 马 莲<sup>2</sup> 颜永红<sup>1</sup>

(1 中国科学院声学研究所语言声学 with 内容理解重点实验室 北京 100190)

(2 北京大学口腔医学院 北京 100081)

(3 北京交通大学信息科学研究所 北京 100044)

2012 年 2 月 7 日收到

2014 年 1 月 8 日定稿

**摘要** 针对腭裂患者易出现塞音弱化或消失的现象, 提出了一种基于塞音段爆破能量检测的腭裂康复手术客观评价方法。该方法采用类听觉的滤波器组作为处理前端, 并对处理后得到的信号在其各子带内分别计算塞音除阻过程中的能量变化率。对腭裂组和术后对照组的平均子带能量变化率进行了对比, 结果表明腭裂组在高频段(子带中心频率从 209.8 Hz 至 8000 Hz) 具有较小的除阻能量变化率。对不送气清塞音 /d/、/b/ 进行了实验, Logistic 回归表明提出的方法与主观判听一致性在音节 /di/ 和 /bu/ 上分别达到 88.9% 和 90.27%。

PACS 数: 43.72, 43.66, 43.80

## Objective evaluation of surgery for cleft palate based on detection of plosive consonants

HU Qi<sup>1,3</sup> ZHAO Qingwei<sup>1</sup> MA Lian<sup>2</sup> YAN Yonghong<sup>1</sup>

(1 Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics  
Chinese Academy of Sciences Beijing 100190)

(2 Peking University School of Stomatology Beijing 100081)

(3 Institute of Information Science, Beijing Jiaotong University Beijing 100044)

Received Feb. 7, 2012

Revised Jan. 8, 2014

**Abstract** By considering the weak form or disappearance of plosive consonants in cleft palate speech, an approach is proposed for automatic diagnosis about if a children's palate is right, or to evaluate if a cleft palate surgery is successful. An auditory filter-bank is used as front-end processor. Then, the energy variation during the release phase of stops within each sub-band is estimated. When the average energy variation within each subband from the cleft palate group and post-palatoplasty group is compared, characteristics of the cleft palate group are identified as decreased energy variation in high frequency bands (209.8 Hz to 8000 Hz). Logistic regression reveals a high consistence (88.9% for syllable /di/, 90.27% for syllable /bu/) between the proposed method and subject evaluation.

## 引言

软腭位于咽腔上部, 在发音时用以阻断鼻腔, 使口音, 特别是爆破音发音完整, 准确。腭裂(cleft

palate)作为口腔颌面部先天畸形(oral cleft)的一种, 在我国发病率为 1.82‰。这种畸形由于口、鼻腔相通, 会导致各种严重的功能障碍, 从而带来一系列社交问题, 对腭裂患者心理造成严重创伤, 影响他们的生活质量。临床上, 语言功能的恢复是腭裂治疗中一

\* 国家自然科学基金(60875014, 60772039)资助项目

† 通讯作者: 赵庆卫, Email: qzhao@hcccl.ioa.ac.cn

个很重要的环节。为了排除患者的语言功能障碍, 首先需要对腭裂患者的语音进行评估, 了解其发音缺陷, 以便进行正确的矫治。

传统的评估方法多采用专家主观判听以及腭咽闭合的形态学观察<sup>[1-2]</sup>。专家对腭裂语音进行评估打分, 易受主观因素影响, 且评分标准难以统一。个人评估结果不可靠, 通常需要对同一样本进行多人分析、多次评估, 工作量比较大。而在形态学分析方面, 需要应用专用设备——鼻咽纤维镜等进行检查<sup>[2]</sup>。虽然能够比较准确, 但诊断费用较高, 不宜大范围的推广。同时, 由于语音是由多个发音器官形态功能的结合产生的, 从而单个发音器官的形态学上观察会有片面性。

正是基于以上考虑, 学者们开始研究腭裂语音的客观评估方法。由于现代语音学<sup>[3]</sup>的发展, 语谱图分析技术受到腭裂语音研究者广泛的关注。诸多研究结果指出, 腭裂患者的元音异常体现为元音不同程度的鼻化<sup>[4-9]</sup>。从语谱图上可以看到共振峰的幅度减小、宽度增加, 出现附加的低频共振峰等。而辅音, 尤其是塞音, 由于构音时间短, 属于暂音, 发音部位和发音方法特殊, 因而发音难度较大, 对于腭裂患者来说, 极易产生发音错误。腭裂患者由于腭部裂开或腭咽闭合不全, 造成发塞音时口腔压力不足, 能量无法集聚, 难以形成正常的瞬间爆破, 在语谱图上就表现为“冲直条 (spike)”或“乱纹 (fills)”的消失或减少<sup>[10-12]</sup>。

语谱图分析技术已经在腭裂语音的研究和腭裂诊断和修复手术评估中有所应用<sup>[5,9-14]</sup>。马思维等<sup>[5]</sup>对不同腭裂类型儿童 /i/ 音鼻音共振峰的特点进行了研究, 提出 /i/ 音鼻音共振峰的频率值以及能量差可以反映不同腭裂类型间的差别。Kataoka<sup>[9]</sup>提出采用 1/3 倍频谱对腭裂儿童异常高鼻化进行检测, 多回归分析表明: 中心频率分别为 1000 Hz, 1600 Hz 和 2500 Hz 三个子带的能量与主观感知打分之间存在高相关 (0.84)。具体而言, 异常高鼻化对应于中心频率分别为 1000 Hz 和 1600 Hz 子带能量的增加以及中心频率为 2500 Hz 子带能量的降低。朱川<sup>[13]</sup>等研究发现, 腭裂语图典型性地表现为: 辅音残缺、共振峰残缺 (第二共振峰时有时无)、贯穿鼻化、清音浊化、噪音干扰、喉塞频现。温波等<sup>[12]</sup>报道了腭裂患者的不送气塞擦音语图模式上缺乏塞音特征, 擦音语图模式能量下降。陈阳等<sup>[10]</sup>通过分析腭裂术后腭咽闭合不全患者辅音的声学特点, 得出如下结论: 腭咽闭合不全患者的语图模式上表现为无冲直条和乱纹或冲直条和乱纹的减少; VOT 缩短或消失; 辅音

脱落或弱化, 清音浊化。朱洪平等<sup>[11]</sup>对 40 名腭裂术前、28 名腭裂术后腭咽闭合不全儿童及 32 名正常儿童的语音声学特征分析表明, 腭裂儿童的塞音和塞擦音的语图上缺乏冲直条、空白间隙等特征, 送气压力性辅音在低频处出现较强噪音能量, 压力性辅音易被省略或浊化。杨增杰等<sup>[14]</sup>利用语谱图上辅音冲直条在不同腭咽闭合状况下的不同表现作为指标, 对腭裂患者术后腭咽闭合状况进行了评价。结果表明, 压力性辅音的冲直条变化与鼻咽镜检查 and 主观判听之间具有良好的一致性。以上研究多是应用语谱图分析对腭裂患者语音特点进行描述, 而应用语谱图分析定量判断患者腭咽闭合功能的方法尚少见。

本文提出了一种客观的定量评估腭裂语音的方法。该方法基于听觉多通道处理, 在听觉域上对塞音信号能量变化率进行定量检测, 并以此为依据对信号中是否存在塞音“脱落”现象进行决策。并且本方法考虑了人耳的听觉特性, 因而显著改善了与主观评分的一致性。

## 1 特征提取

辅音的发音一般包括成阻、持阻和除阻三个阶段。压力性辅音中的塞音和塞擦音会同时具有这三个阶段。除阻是声道阻塞的突然打开, 声道压力的骤然释放。在语谱图上表现为一个持续时间很短, 频带分布很宽的“冲直条”。图 1 示出了音节 /di/ 的时域波形及其语谱图形式, 并标出了塞音 /d/ 的爆破。

本文提出了一种基于听觉滤波器组的多通道分析方法。该方法利用了“冲直条”宽频域和窄时域的

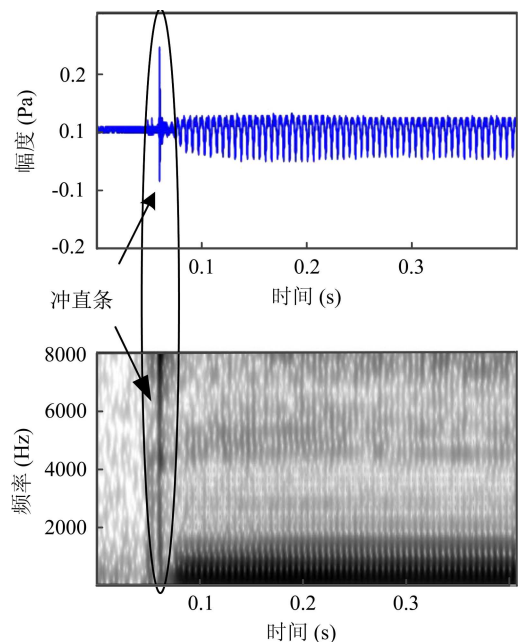


图 1 音节 /di/ 的时域波形及其宽带语谱图

特性。在宽频域方面, 利用听觉多通道处理将时域上的原单脉冲转换沿子带分布的多脉冲。然后分别计算各子带内脉冲特征, 以构成一多维特征向量。在窄时域方面, 考虑到加窗处理会影响“冲直条”的短时特性, 采用以采样点为基础的累加能量曲线进行爆破检测, 并提取相应特征进行腭裂语音的评价。

### 1.1 听觉谱域信号分解

为了在听觉谱域提取信号特征, 首先将信号输入一听觉滤波器组, 以实现听觉子带分解。

#### 1.1.1 外耳和中耳模型

人的外耳和中耳会对在 2 ~ 4 kHz 范围内的声音信号进行放大, 该信号增益改变可以通过线性内插等响度曲线获得, 然后将增益改变加权至后续的听觉滤波器组中。

#### 1.1.2 听觉滤波器组

听觉滤波器组的设计, 采用 Patterson<sup>[15-16]</sup> 提出的 Gammatone 滤波器组来模拟耳蜗的频率选择特性<sup>[16]</sup>。图 2 为 80 ~ 5000 Hz 频率范围内 20 个 Gammatone 滤波器的频率响应图。从图中可以看出, 滤波器的中心沿对数频率轴呈等间隔分布, 而且滤波器带宽随着中心频率的增加而增大。本文采用 128 个听觉滤波器用以覆盖语音信号 50 ~ 8000 Hz 的频率范围, 该设置足以实现计算复杂度和频率分辨率之间的折衷<sup>[16]</sup>。

Gammatone 滤波器的脉冲响应为<sup>[15-16]</sup>:

$$g(n, f_c) = (n/f_s)^{N-1} \exp[-2\pi b(f_c)] \cos(2\pi f_c n/f_s + \varphi) u(n/f_s), \quad (1)$$

其中  $N$  是模型阶数,  $f_c$  是第  $c$  个滤波器的中心频率 (Hz),  $\varphi$  是初始相位,  $u(n/f_s)$  是单位阶跃函数,  $f_s$  是

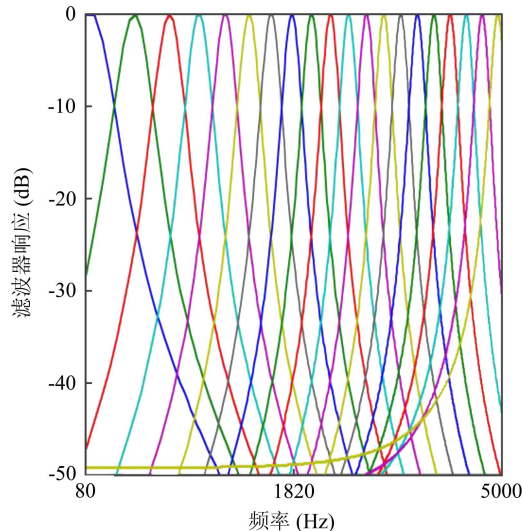


图 2 20 子带 Gammatone 滤波器组响应示意图

信号的采样频率,  $b(f_c)$  是对应中心频率为  $f_c$  时滤波器的带宽。一般取  $N = 4$ , 而  $b(f_c) = 1.019\text{ERB}(f_c)$ ,  $\text{ERB}(f_c) = 24.7 + 0.108f_c$ 。其中, ERB 是 Equivalent Rectangular Bandwidth 的缩写<sup>[16-17]</sup>。

假定  $s(n)$  为输入信号, 对每一个滤波器通道  $c$ , 其相应滤波后的信号为:

$$s(n, c) = g(n, f_c) * s(n), \quad 1 \leq c \leq 128 \quad (2)$$

其中“\*”表示卷积运算。由于低频滤波器频率带宽较小, 因而通过它滤波得到的信号包络峰值会存在一个较大的延时; 而高频滤波器与此正好相反。因而需要对滤波后的信号进行延时, 以补偿滤波器带来的延时。Patterson<sup>[15-16]</sup> 提出将各通道输出后的信号向前延时  $t_c = (N - 1)/2\pi b(f_c)$  以对齐各通道的包络峰值。

### 1.2 冲直条能量分析

信号经 Gammatone 滤波器组处理以后, 得到的是原信号在不同频率尺度下的时域信息。对该时域信号以采样点为单位进行能量累加处理, 以得到直观反映塞音冲直条 (爆破段) 的能量突变特征<sup>[18-19]</sup>。

设  $s(n, c)$  是第  $c$  个 Gammatone 滤波器的输出信号,  $E_c(n)$  是该信号在第  $n$  时刻的累计能量, 它们之间满足如下关系式:

$$E_c(n) = E_c(n - 1) + s(n, c) s(n, c), \quad (3)$$

图 3 和图 4 分别示出了正常塞音样本和腭裂塞音样本在各个子带内的累计能量。为了清晰起见, 图中只绘出了第 1, 21, 41 子带内的能量曲线图, 其对应塞音样本的时域波形也绘于该图的底部。从图中可以看出, 正常塞音样本在爆破发生时刻会出现能量

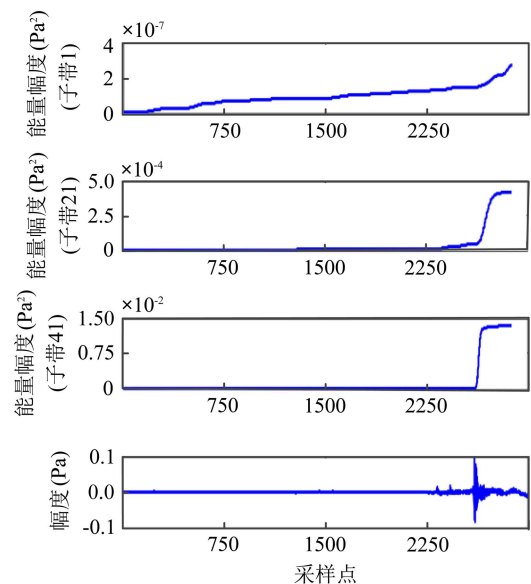


图 3 正常塞音样本多尺度处理后的能量曲线图

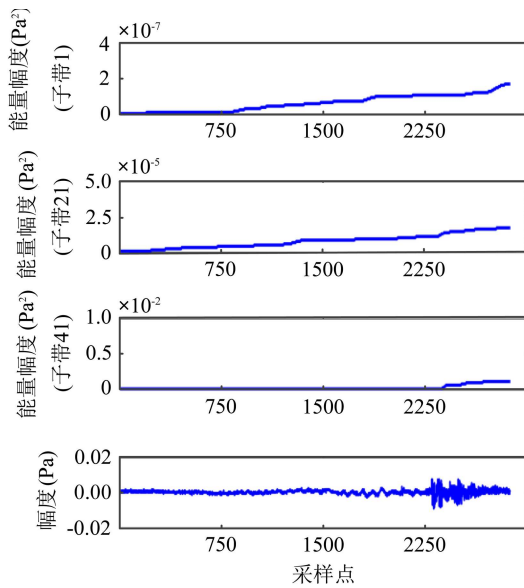


图 4 腭裂塞音样本多尺度处理后的能量曲线图

曲线的快速上升, 而腭裂语音由于缺乏或弱化爆破阶段, 在整个时域上能量曲线只表现为稳中有升, 其能量变化远没有正常塞音快。从图 3 和图 4 中还可以看出, 这种能量变化的差异是与具体子带相关联的。具体而言, 两种塞音在第 1 子带内的能量曲线差异较小, 而在第 21 和 42 子带内的差异却较大。因此, 本文选取塞音除阻时, 能量曲线的变化率作为爆破检测的特征。

### 1.3 能量变化率计算

为了计算能量变化率, 需要找到对应于爆破段的起止点。由于紧临爆破段前后, 能量的变化都比较平缓, 因而可以通过比较累计能量曲线的变化率, 找到爆破段的起止点。

具体过程如下:

(1) 在各子带内, 计算能量曲线的差分, 即:  $D_c(n) = E_c(n) - E_c(n-1)$ 。其中,  $E_c(n)$  表示第  $c$  个子带内的累加能量曲线, 其由式 (3) 计算得到,  $c \in \{1, 2, 3, \dots, 128\}$ 。

(2) 查找  $D_c(n)$  的极大值, 并确定其所对应的时刻点  $k_{c\_end}$ , 并将该点定为第  $c$  子带爆破段的结束点。

(3) 以该点为起始点, 与其前向各点作连线, 并求各连线的斜率:

$$\mu_c(k) = \frac{1}{k_{c\_end} - k} (E_c(k_{c\_end}) - E_c(k)), \quad (4)$$

$$(0 \leq k \leq k_{c\_end} - 1)$$

将斜率变化最大的点作为爆破段的起始点  $k_{c\_start}$ , 而此时的最大斜率值就是所要求的能量变化率  $\sigma_c$ 。图 5 给出了该算法应用于正常样本的检测结果 (第 110 子带)。其中, 从上至下分别表示子带能量曲线

$E_{110}(n)$ , 子带能量曲线差分函数  $D_{110}(n)$  和斜率变化曲线  $\mu_{110}(k)$ 。

现将上述算法流程总结如下:

(1) 输入信号进行听觉域多通道分析, 并得到相应子带的时间信号;

(2) 各子带时域信号按式 (3) 进行时域能量累加, 以计算各子带内能量变化曲线  $E_c(n)$ ;

(3) 在各子带内分别查找爆破段的起止点, 并按照式 (4) 计算对应爆破段的能量变化率  $\sigma_c$ 。

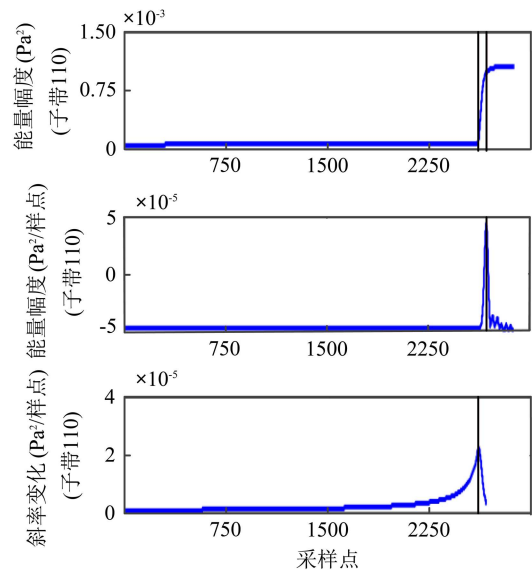


图 5 爆破段检测示意图

## 2 实验

### 2.1 样本选择

#### 2.1.1 研究对象

选取来北京大学口腔医学院唇腭裂治疗中心就诊的腭裂儿童和腭裂术后正常儿童共 48 名。其中男 29 名, 女 19 名。(1) 腭裂组: 腭裂儿童 31 名, 男 18 名, 女 13 名。年龄分布从 6—12 周岁 (年龄 7 至 11 岁, 平均 9.75 岁) 和 12 周岁以上 (年龄 13 至 47 岁, 平均 21.57 岁) 两组。所有参与者均无明显智力、听力缺陷, 无重度牙颌畸形、上呼吸道感染且普通话熟练。(2) 对照组: 术后正常儿童 17 例, 年龄与患儿年龄相匹配, 男 11 例, 女 6 例; 所有参与者均智力、语言发育正常, 无牙颌畸形。

#### 2.1.2 录音设备

(1) Praat 语音分析软件; (2) 硬件: Sony IC Recorder Icd-UX80 录音笔及卡式麦克风。

#### 2.1.3 语音采集

录音前向参加实验的儿童简述实验步骤, 并得到配合。录音环境安静, 实验儿童读 /di/、/bu/ 各

三遍。

### 2.1.4 语音材料

所有语料经 3 位长期从事腭裂语音诊治工作的临床医师判听，当 3 人判听均认为发音异常或正常才将其纳入研究范围。用于分析的样本共 288 例 (/di/、/bu/ 各 144 例)，其具体信息如下：腭裂组 93 例 (男 54 例，女 39 例)，术后对照组 51 例 (男 33 例，女 18 例)。所有样例采样率均为 16 kHz，动态范围为 60 dB。所有统计分析均采用 SPSS 11.5 软件。

### 2.2 实验结果

为了进行比较，图 6 示出了塞音 /d/ 和 /b/ 的平均子带能量变化率曲线。图中所有样本第一子带能量变化率调整为 -6 dB。所有子带能量变化率都采用第一子带能量变化率进行归一化。通过比较两组的平均多带能量变化率曲线可以看出：(1) 对照组样本在高频处 (233.6 Hz 以上) 的能量变化率比腭裂组有明显增加，相较之下，低频处 (50 ~ 233.6 Hz) 两组间能量变化率差异较小。而且，高频处的差异比较稳定，未出现随子带差异的大起大落现象。(2) 塞音 /d/ 和 /b/ 差异相似。

为进一步定量分析各子带差异，采用 *t* 检验对各

子带内两组间均值的差异进行显著性比对。结果表明：对于 /di/ 和 /bu/，从第 18(中心频率 209.8 Hz) 至 128 子带 (中心频率为 8000 Hz) 均具有统计显著性差异 ( $P < 0.01$ )。

依据以上分析，后文只选取具有显著区分性的子带能量变化率系数进行分析，而对不具有区分性的子带不再予以考虑。由于各子带能量变化率之间存在高度相关 (以 /di/ 为例，其第 127 和 128 子带能量变化率的相关系数为 0.98,  $P < 0.001$ )，因而采用主成分分析<sup>[20]</sup>对得到的多维子带特征进行正交化处理。得到的前 5 个特征值及对应和方差贡献率示于表 1。从表 1 中可以看出，前 5 个特征值的累计方差贡献率大于 90%。

经过主成分分析后，原多个子带特征都被变换为 5 个相互正交的特征。对 /di/ 和 /bu/，分别构建 Logistic 回归模型以分析 5 个特征与主观判听的关系。Logistic 回归模型具有如下形式：

$$y = 1/(1 + \exp(-z)), \tag{5}$$

其中， $z = \alpha + \sum_{i=1}^N \beta_i M_i$ ， $\alpha, \beta_i (i = 1, \dots, N)$  为待求参数； $M_i (i = 1, \dots, N)$  表示测量得到的声学特征， $N$  表示声学特征数量； $y$  表示输入样例为腭裂的概率。

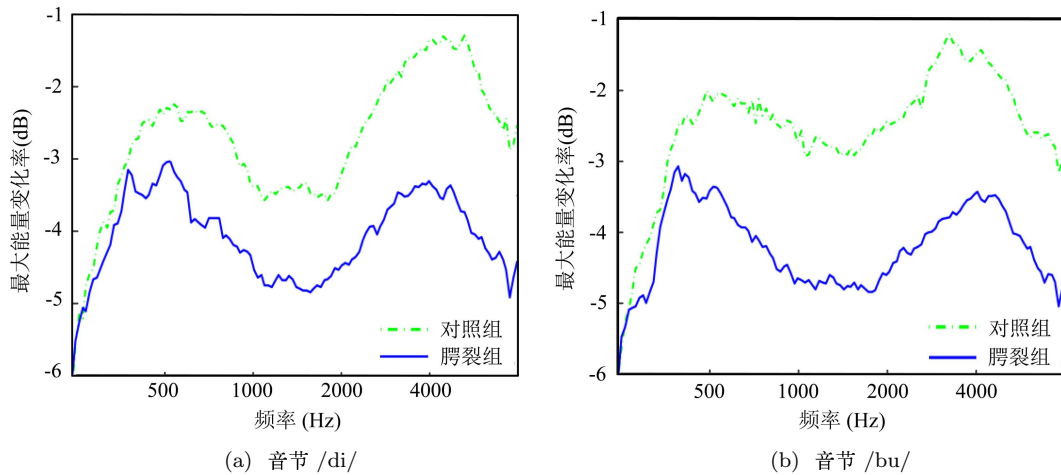


图 6 腭裂和术后对照组平均子带能量变化率示意图

表 1 特征值及对应方差贡献率

	/di/		/bu/	
	特征值	方差贡献率 (%)	特征值	方差贡献率 (%)
成分 1	9154	76.84	7864.1	81.02
成分 2	773.774	6.49	496.68	5.1
成分 3	338.2	2.84	296.14	3.05
成分 4	291.62	2.45	120.7	1.25
成分 5	206.5	1.73	80.96	1.16

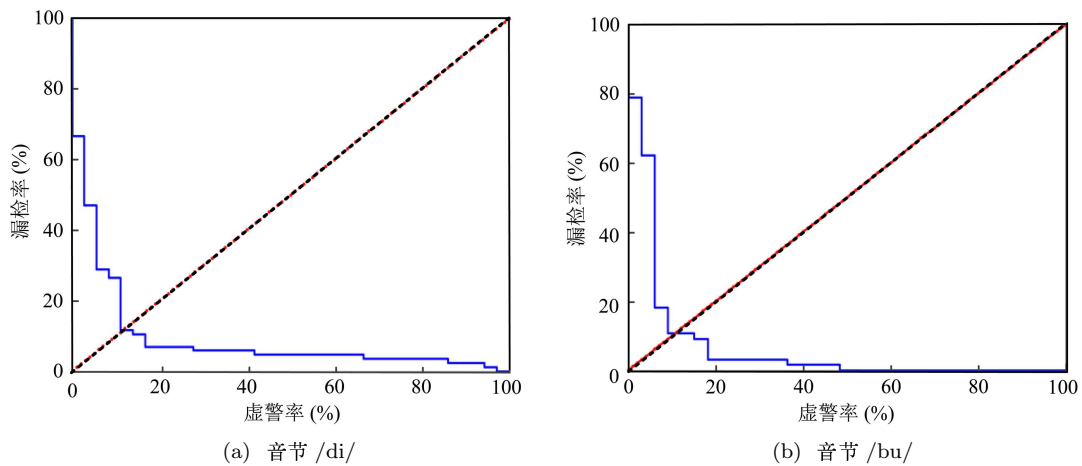


图 7 算法的 DET 曲线

表 2 分类表

	预测值 *		百分比校正 (%)	预测值 **		百分比校正 (%)
	/di/			/bu/		
观测值	对照组	腭裂组		对照组	腭裂组	
对照组	45	6	88.2	46	5	90.2
腭裂组	10	83	89.2	9	84	90.3
总计百分比			88.9			90.27

\* 切割值为 0.7, \*\* 切割值为 0.67

对于 /di/ 和 /bu/, 得到的回归方程分别为:

$$y_{di} = 1 / (1 + \exp(-0.274 * P1 + 8.764)), \quad (6)$$

$$y_{bu} = 1 / (1 + \exp(-0.319 * P1 + 10.972)), \quad (7)$$

其中,  $P1$  是经主成分分析得到的声学参数 (对应于最大特征值);  $y$  为预测值, 最后的分类结果采用  $\delta$  的切割值。即: 当样例预测值  $y > \delta$  时, 该样例被判决为腭裂。反之, 则判决为正常。注意到,  $P2, P3, P4, P5$  等声学参数并没有出现在回归方程中, 这是因为这些声学参数在对样例判决中并不具有显著性影响 ( $P > 0.05$ )。

为了确定  $\delta$  的取值, 图 7 绘出了本文方法对音节 /di/ 和 /bu/ 的错误检测折中曲线<sup>[21]</sup> (Detection Error Tradeoff, DET)。本文选取虚警率和漏检率相等时的切割值  $\delta_0$  作为待求目标。从图中可以看出, 对于 /di/, 虚警率和漏检率约为 11%(对应切割值  $\delta_0$  为 0.7); 对于 /bu/, 此时虚警率和漏检率约为 10%(对应切割值  $\delta_0$  为 0.67)。

根据选取的切割值进行 Logistic 回归, 并将得到的分类结果与主观判听结果进行比对, 并计算相应正确百分比 (表 2)。表 2 中, “观测值” 为主观判听结果, “预测值” 是由式 (6) 和式 (7) 并结合求得的切割值计算得出。

从表中可以看出, 对于 /di/, Logistic 函数全局

分类正确率达到 88.9%。其中, 虚警率 (样本主观判听为正常, 而算法判定为异常) 为 11.8%; 漏检率 (样本主观判听为异常, 而算法判定为正常) 为 10.8%; 对于 /bu/, 全局分类正确率达到 90.27%。其中, 虚警率为 9.8%; 漏检率为 9.7%。

### 2.3 讨论

上文已经指出, 无论对于 /di/ 还是 /bu/, 正常和腭裂塞音的差异主要出现在高频段 (209.8 Hz ~ 8000 Hz); 而且, 该差异相对稳定, 并不存在随子带不同而产生的巨大波动。同时, Logistic 回归表明, 正交化后的单一主成分与主观判听具有高度一致性, 这预示着各子带的能量变化率参数之间存在高度相关。在实际操作中, 可以将本文算法进行简化: 设计一个通带起始频率为 210 Hz 的高通滤波器, 并利用该滤波器对待评价样本进行处理, 然后对处理得到的信号计算累计能量变化率, 并以此作为腭裂爆破语音评价的客观标准。

### 3 结论

本文提出了一种腭裂语音客观定量的评估方法。该方法采用类听觉滤波器组对信号进行时频处理, 并对处理后的各听觉子带进行塞音能量变化率的检测。对腭裂组和术后对照组的平均子带能量变

化率采用  $t$  检验进行对比, 结果表明腭裂组在高频段 (子带中心频率从 209.8 Hz 至 8000 Hz) 具有较小的除阻能量变化率 ( $P < 0.01$ )。利用主成分分析对选取的子带特征进行正交化, 并采用 Logistic 回归进行分类预测。在音节 /di/ 和 /bu/ 上的实验表明, 提出的方法与主观聆听的一致性分别达到 88.9% 和 90.27%。作为辅助特征, 该方法能够与采用元音特征<sup>[4-5,12,22]</sup> (如: 共振峰频率和带宽, 虚拟共振峰检测) 的方法相结合, 以形成完整的腭裂语音客观评估系统。

### 参 考 文 献

- 1 Kuehn D P, Moller K T. Speech and language issues in the cleft palate population: the state of the art. *The Cleft Palate-Craniofacial Journal*, 2000; **37**(3): 348.1—348.35
- 2 朱敏, 袁文化, 卢晓峰等. 应用鼻咽喉纤维镜评价腭咽闭合功能的定量分析研究. *中华口腔医学杂志*, 1998; **33**(3): 172—174
- 3 吴宗济, 林茂灿. 实验语音学概要. 北京: 高等教育出版社, 1989
- 4 封兴华, 魏建华, 王兵等. 腭裂术后患者语音频谱分析. *第四军医大学学报*, 2001; **22**(19): 1794—1796
- 5 马思维, 文抑西. 不同腭裂类型儿童 /i/ 音鼻音共振峰分析研究. *中国美容医学*, 2011; **20**(1): 48—53
- 6 Kent R D, Read C. The acoustic analysis of speech. 2nd edition, California: Singular Publishing Group, 2001: 211
- 7 Spriestersbach D C, Darley F L, Rouse V. Articulation of a group of children with cleft lips and palate. *Journal of Speech and Hearing Research*, 1956; **21**(4): 436—445
- 8 Gibbon F, Smeaton-Ewins P, Crampin L. Tongue-palate contact during selected vowels in children with cleft palate. *Folia Phoniatri Logop*, 2005; **57**(4): 181—192
- 9 Kataoka R, Warren D W, Zajac D J *et al.* The relationship between spectral characteristics and perceived hypernasality in children. *Journal of Acoustical Society of America*, 2001; **109**(5): 2181—2189
- 10 陈阳, 杨育生, 吴忆来等. 腭裂术后腭咽闭合功能不全患者辅音声学特点的研究. *中国口腔颌面外科杂志*, 2005; **3**(4): 279—284
- 11 朱洪平, 孙勇刚, 王光和等. 腭裂儿童的辅音声学特点研究. *现代口腔医学杂志*, 1999; **12**(3): 181—183
- 12 温波, 张学军, 吕明臣等. 腭裂语音中不送气辅音的声学特征. *吉林大学学报 (医学版)*, 2004; **30**(5): 769—771
- 13 朱川, 李士忠. 腭裂语音矫治实验研究. *语言文字应用*, 2000(1): 77—81
- 14 杨增杰, 陈仁吉, 穆口. 应用语图分析评价腭裂术后腭咽闭合状况的临床研究. *北京口腔医学*, 2010; **18**(1): 36—38
- 15 Patterson R, Nimmo-Smith I, Holdsworth J *et al.* An efficient auditory filter-bank based on the gammatone function. *Applied Physics*, 1988
- 16 Wang D L, Brown G J *et al.* Computational auditory scene analysis: principles, algorithms, and applications. IEEE-Wiley Press, 2006: 15—18
- 17 Huang X D, Acero A, Hon H W. Spoken language processing: a guide to theory, algorithm and system development. Prentice Hall, 2001: 30—34
- 18 董滨. 计算机辅助汉语普通话学习和客观测试方法的研究. 博士学位论文, 中国科学院声学研究所, 2006
- 19 Du L M, Hou Z Q. Automatic extraction of stop-oriented features from Chinese speech wave using wavelet transform. In Proceedings of ICSP' 96, Beijing, 1996(1): 303—306
- 20 Duda R O, Hart P E, Stork D G. Pattern classification. 2nd edition, Wiley-Interscience 2000: 568
- 21 Martin A, Doddington G, Kamm T *et al.* The DET curve in assessment of detection task performance. Proceedings of Eurospeech'97, Rhodes, Greece, 1997(4): 1895—1898
- 22 封兴华, 魏建华, 张立军等. 频谱分析技术在腭裂语音研究中的应用. *中国口腔颌面外科杂志*, 2003; **1**(2): 70—73