

一种实现耳鸣匹配声源多样化的新方法

熊培杰 何培宇 潘 帆

(四川大学电子信息学院, 成都 610065)

摘 要: 在耳鸣掩蔽治疗领域中, 由于患者对同一种耳鸣匹配声源的感受千差万别, 因此其多样性是影响治疗效果的决定性因素。但国内外传统的治疗仪普遍存在耳鸣匹配声源不足的缺陷。针对这个问题, 本文受语音转换系统的启发, 提出了一种基于双线性变换函数的方法来对特定的匹配声源进行频谱搬移以达到声源多样化的目的。同时, 由于双线性变换函数只能有效改变声源共振峰的频率以及带宽而不能对其幅度进行控制, 这样的特性最终会导致其能量的衰减。那么文中提出利用临界倾斜滤波器来补偿频谱搬移所引起的能量损失, 并给出了一种搜寻最佳倾斜系数的方法及其微调范围。通过仿真表明, 这种方法取得了良好的效果, 丰富了耳鸣匹配声源的多样性, 具有一定的实用参考价值。

关键词: 双线性变换函数; 临界倾斜滤波器; 频谱搬移

中图分类号: TN912 **文献标识码:** A **文章编号:** 1003-0530(2011)08-1259-06

A New Method to Realize the Variety of Tinnitus Matching Sound

XIONG Pei-jie HE Pei-yu PAN Fan

(School of Electronics and Information, Sichuan University, Chengdu 610065, China)

Abstract: In the field of tinnitus masking treatment, patients have different feelings of the same tinnitus matching sound, so the variety of sound sources is an important determinant of treatment effect. However, the lack of tinnitus matching sound sources ubiquitously exists in traditional treatment equipments. Thus, inspired by voice conversion system, a method of spectrum shifting which based on bilinear transform function is proposed to realize the variety of tinnitus matching sound. Meanwhile, bilinear transformation function can only effectively alter the resonant frequency and bandwidth but can not control its amplitude, so this kind of character will finally cause attenuation to energy, in order to compensate the energy attenuation the critical tilt filter is used and a new method of finding the optimum tilt factor is proposed. In addition, the range of optimum tilt factor is also discussed. Simulation results show that this method has achieved good effect and enriched the variety of tinnitus matching sound, and it has certain practical reference value.

Key words: bilinear transform function; critical tilt filter; spectrum shifting

1 引言

耳鸣(Tinnitus)是指在无外源性的声或电刺激情况下,人耳产生的一种声音感觉。常被形容为嘶嘶声、嗡嗡声等。其发病机理复杂,属耳科三大难症(耳鸣、耳聋、眩晕)之一^[1]。

目前较为行之有效的治疗手段为掩蔽疗法,其原理是用与耳鸣匹配的声刺激产生掩蔽效应,以促进对耳鸣的适应。国际上较为成熟的治疗设备是加拿大一实验室开发的听尼特(TinniTest)耳鸣综合诊断治疗仪,它通过增加匹配声源的种类来丰富匹配声源的多样性,但存在不能对特定声源(如蝉鸣声)进行频率调节的问题^[2,3]。而对于不同的患者来讲,他们对同一种耳鸣匹配声音的感受是存在明显差异的,这样就大大降低了对不同病人的匹配度。相较国际上的掩蔽治疗

仪,国内大多数医院采用的治疗设备在功能上更加局限,其声源种类不但稀少并且对特定声源不能进行频率调节,匹配效果非常不理想。但通过查新发现,国内外均未见有相关的研究报告来改善临床上耳鸣掩蔽疗法中掩蔽声单一、匹配度低、可调性差的问题。那么可以借鉴在人类语音转换中已趋于成熟的算法来尝试实现医用掩蔽声的多样性,从而填补这一片空白。

在语音转换领域中,双线性变换函数实现谱包络变换所需的参数很少,只要一个参数就可对共振峰的特性进行改变。相较其他方法,如重采样法、GMM法、HMM法等其在算法简单易行,能够满足软件工程中的实时性处理要求。并且GMM法、HMM法需要提取源语音与目标语音的相关参数来达到语音转换的目的。但这里我们没有目标语音,仅是通过调节参数来合成语音,因此从这两方面考虑,提出了一种利用双线性

性变换函数以及临界倾斜滤波器相结合的新方式来实现特定匹配声源的多样性,从而对当今耳鸣诊疗设备所存在的缺陷起到一定的改善作用。

2 频谱搬移系统

一般地,我们可以利用激励信号激励一个线性时不变系统来模拟产生语音的过程^[4]。假设系统为全极点模型,首先通过线性预测求得LPC系数和残差信号,然后对LPC系数进行双线性变换,最后利用残差信号激励频移后的系统就可得到合成语音^[5,9]。

借用这样的思想,我们将双线性变换函数用于耳鸣常见声音的频谱搬移,来达到改变其共振峰频率以及带宽的目的,从而将某一种耳鸣匹配声音变得更加多样。同时由于患者的耳鸣声如蝉鸣声、水流声、蚊子声为持续、稳定、单一的声音,在掩蔽治疗阶段可以近似认为其具有长时平稳特性。因此文中所有实验均以0.2秒(约10000点)为一帧进行处理,并最终利用合成的时间长度为2秒的声音通过循环播放的方式来实现耳鸣掩蔽治疗。除此之外还利用临界倾斜滤波器来弥补双线性变换函数不能控制共振峰幅度的缺陷。

系统框图如下所示:

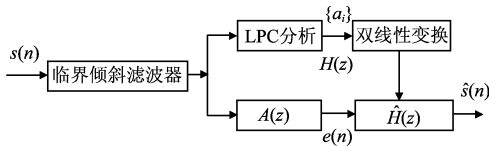


图1 频谱搬移系统
Fig. 1 Spectrum shifting system

图1中 $s(n)$ 为耳鸣匹配音源, $e(n)$ 为残差信号, $\{a_i\}$ 是LPC系数, $H(z)$ 为声道传输函数, $\hat{H}(z)$ 为频谱搬移后的声道传输函数, $\hat{s}(n)$ 为最终合成的耳鸣匹配声音。临界倾斜滤波器在这里主要起到保护耳鸣匹配音源能量的作用,将在2.2节作详细介绍。

声道传输函数 $H(z)$ 的表达式为:

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}} \quad (1)$$

其中LPC系数 $\{a_i\}$ 可利用求解自相关方程的Durbin算法获得。

2.1 双线性变换函数

双线性变换函数为:

$$z = \frac{\beta + \hat{z}}{1 + \beta \hat{z}} \quad (2)$$

双线性转换函数的频率弯折关系为:

$$\hat{\omega} = \arctan\left(\frac{(1-\beta^2)\sin\omega}{(1+\beta^2)\cos\omega - 2\beta}\right) \quad (3)$$

上式中 β 为频率弯折系数。当 $\beta > 0$ 时,频谱被移向高频处;当 $\beta < 0$ 时,频谱被移向低频处。当且仅当 $|\beta| < 1$ 时可保证系统稳定,其推导过程见文献^[6]。

实验中我们利用上述中的双线性变换法对耳鸣中比较常见的声音(蝉鸣声)进行频谱搬移,其发声模型的幅频特性如下图所示(未采用临界倾斜滤波器进行补偿):

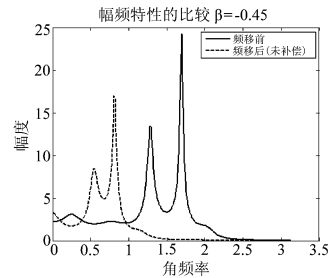


图2 幅频特性 $\beta = -0.45$

Fig. 2 Amplitude-frequency characteristic($\beta = -0.45$)

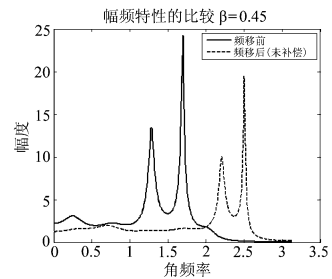


图3 幅频特性 $\beta = 0.45$

Fig. 3 Amplitude-frequency characteristic($\beta = 0.45$)

从图2与图3中可以看出,当 β 分别取 -0.45 和 0.45 的时候,我们可以得到如下的三条信息:1、频谱搬移后蝉鸣声共振峰的位置得到了明显的搬移。2、频谱搬移后蝉鸣声共振峰的带宽发生了明显的变化。3、频谱搬移后蝉鸣声共振峰的幅度有明显的降低。这几点也正是实现蝉鸣声多样性的关键因素。但基于双线性变换函数的特性,它只能合理地掌控共振峰频率以及带宽的变化,并不能对其幅度进行有效的控制。对蝉鸣声来说,随着频谱搬移程度的加大,共振峰的幅度会越来越小,最终导致蝉鸣声能量的衰减。

由于篇幅有限,这里仅通过对比当 $\beta = -0.45$ 时蝉鸣声的时域波形图与原始蝉鸣声的时域波形图来验证上述中能量衰减的存在。采样点数为10000个点(约0.2秒),实验结果如图4、图5所示:

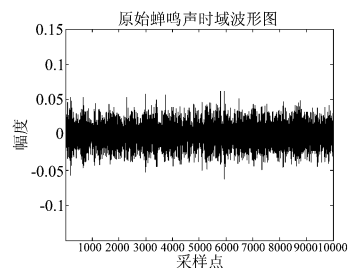


图4 原始蝉鸣声时域波形

Fig. 4 Time domain waveform of cicadas sound(original)

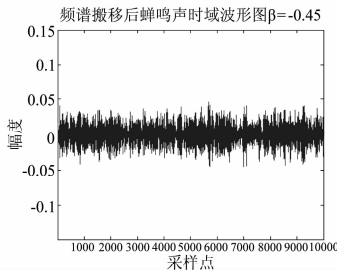


图 5 搬移后蝉鸣声时域波形

Fig. 5 Time domain waveform of cicadas sound(after spectrum shifting)

从图 5 中可以明显地发现,当我们用双线性变换函数对蝉鸣声进行频谱搬移后($\beta = -0.45$),其时域波形图整体上是小于原始时域波形图的,也就是说其能量有所衰减。

由于蝉鸣声第一共振峰和第二共振峰的幅度也是影响其多样性的因素之一,那么在有效弥补能量衰减的基础上非线性的控制两个共振峰幅度之间的对应关系显然要比线性的去控制它们有用得多,因此引入临界倾斜滤波器是一个好的选择。

2.2 临界倾斜滤波器

频谱倾斜是语音信号中一个很重要的语音个性特征^[6],主要体现在频谱搬移前后不同语音信号中共振峰的带宽和幅度有所不同(如图 4 所示)。在对蝉鸣声进行频谱搬移的过程中,其两个共振峰的幅度会出现明显的降低,最终导致能量的衰减。一般地,通过引用临界倾斜滤波器相关的两个函数($H(z)$ 、 $Q(z)$)来对能量损失进行非线性补偿。

临界倾斜滤波器函数^[7]:

$$Q(z) = (1 + g * z^{-1})^2 = 1 + 2 * g * z^{-1} + g^2 * z^{-2} \quad (4)$$

$$H(z) = \frac{1}{Q(z)} = \frac{1}{1 + 2 * g * z^{-1} + g^2 * z^{-2}} \quad (5)$$

式(4)、(5)中的 g 为倾斜系数, g 的值是使式(6)达到最小时所对应的值:

$$c_0 g^3 + 3c_1 g^2 + (2c_0 + c_2)g + c_1 = 0 \quad (6)$$

其中 $c_i = \sum_{n=i}^{N-1} s(n)s(n+i)$ 为短时自相关函数, g 的取值范围应在 $[-1 \sim +1]$ 之间^[7]。

在对蝉鸣声等音频信号的补偿过程中,与对语音信号进行补偿不同的是,这里没有既定的目标音源 $y(n)$,因此这里通过计算频谱搬移前的倾斜系数 g 与频谱搬移后的倾斜系数 g_s 的关系来选择该用 $H(z)$ 函数还是 $Q(z)$ 函数进行补偿。一般地,当 $|g_s| < |g|$ 选择 $Q(z)$ 函数来补偿,当 $|g_s| > |g|$ 选择 $H(z)$ 函数来补偿^[6]。结合上一章频谱搬移的步骤来实现能量补偿的具体过程如下:

首先利用双线性变换函数对耳鸣匹配声源 $s(n)$ 进行频谱搬移得到 $\hat{s}(n)$,再利用(6)分别求得频移前的倾斜系数 g 与频移后的倾斜系数 g_s ,并比较它们之间

的关系。如果 $|g_s| < |g|$ 那么选择 $Q(z)$ 函数对耳鸣匹配声源进行补偿,否则选择 $H(z)$ 函数对耳鸣匹配声源进行补偿。最后利用式(7)来确定最佳 g 值。

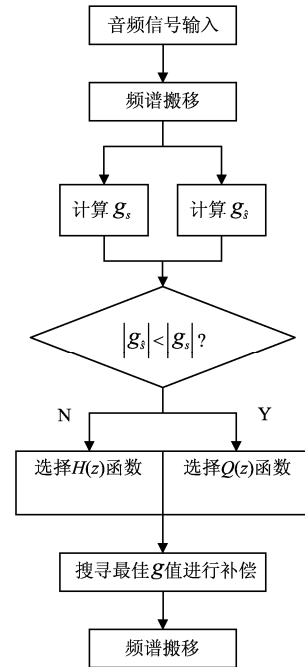


图 6 能量补偿流程图

Fig. 6 Flow chart of energy attenuation

$$Per(g) = \left(\sum_{k=1}^n (E_s(k, g))^2 / \sum_{k=1}^n (E_s(k, g_s))^2 \right) * 100\% \quad (7)$$

其中 $E_s(k, g)$ 对应频谱搬移前第 k 点在补偿系数为 g 时的幅度, $E_s(k, g_s)$ 对应频谱搬移后第 k 点在补偿系数为 g_s 时的幅度, n 为一帧声音信号的点数, g 的取值区间为 $[-1 \sim +1]$ 。当 g 的取值使得 $Per(g)$ 取最大值的时候即为最佳补偿系数(需要注意的是有可能出现 $Per(g) > 100\%$ 的情况,说明补偿过度了,应约束 $Per(g) = 100\%$ 即达到最佳)。

3 仿真实验

3.1 对蝉鸣声源的仿真

接下来通过仿真实验来说明临界倾斜滤波器在实现蝉鸣声多样性这个系统中所发挥的作用。实验中,仍然使用蝉鸣声作为我们的主要研究对象,令 $\beta = 0.25$,采样点数取 10000 点。

图 7 中实线为蝉鸣声的原始幅频特性、带圈实线为蝉鸣声频谱搬移后(未补偿)的幅频特性,带三角实线为蝉鸣声频谱搬移后(补偿)的幅频特性。从图中可以明显看出,经过补偿后,蝉鸣声第二共振峰的幅度有显著的提升,但第一共振峰的幅度几乎保持了不变,也就是说通过临界倾斜滤波器对两个共振峰幅度的对应关系进行了有效的控制。

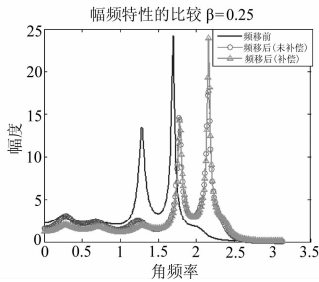


图7 幅频特性的比较

Fig. 7 Compare of amplitude-frequency characteristic

图8中的三幅图从上到下分别为原始蝉鸣声的时域波形图、蝉鸣声频谱搬移后(未补偿)的时域波形图、蝉鸣声频谱搬移后(补偿)的时域波形图。从图中可以看出,补偿后蝉鸣声的时域波形图整体上已与原始蝉鸣声的时域波形图比较相近。

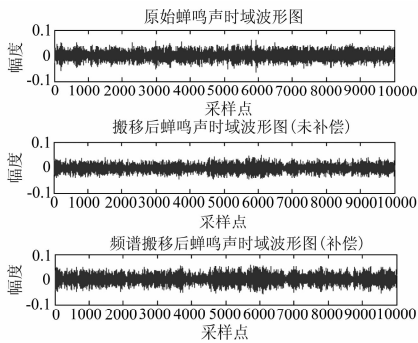
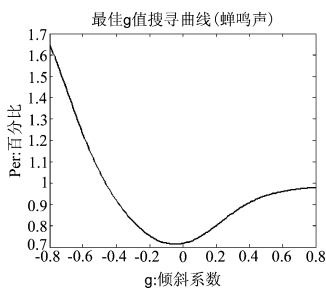


图8 时域波形图的比较

Fig. 8 Compare of time domain waveform

那么结合图7和图8的实验可以看出,临界倾斜滤波器在弥补能量衰减的同时也对蝉鸣声第一共振峰和第二共振峰的对应关系进行了有效的控制,达到了预期的效果。

图9是与图7相对应的最佳 g 值搜寻曲线,从图中不难看出,当 $g=-0.5$ 左右时即达到了比较好的补偿效果(能量相似度接近100%, $g=0$ 为没有补偿的情况)。

图9 最佳 g 值搜寻曲线Fig. 9 Best g value search curve

特别地,由于蝉鸣声共振峰幅度的大小也是决定其声音多样性的因素之一。虽然 g 存在一个最佳补偿

系数,但是在实际应用中,由于不同耳鸣患者的感受千差万别,在补偿过程中,稍微过补偿一点或者欠补偿一点都是可以的,反应在蝉鸣声的幅频特性上也就是第一和第二共振峰幅度的对应关系会产生明显的不同,那么合理对 g 取进行取值反而会起到让蝉鸣声变得更加丰富的作用。这样的特性也是采用直接频移法所不具备的,因为直接频移法无法对共振峰的幅度进行有效控制。(建议以最佳系数为中心,前后进行最大0.1左右差值的微调,调整幅度过大有可能使得能量的补偿效果不佳)

由于篇幅有限,表1为蝉鸣声在不同 g 值情况下(实际取值不超过最佳系数的正负0.1)补偿前后能量百分比的统计,试验条件 $\beta=0.25$,采样点数仍取10000个点,实验结果如表1所示:

表1 不同 g 值能量百分比的统计Tab. 1 Statics of energy percentage using different g value

音频源	最佳系数 g 值	实际系数 g	补偿前 $Per(g)$	补偿后 $Per(\hat{g})$
蝉鸣声	-0.5	-0.56	85.4%	115.3%
蝉鸣声	-0.5	-0.54	85.4%	108.7%
蝉鸣声	-0.5	-0.52	85.4%	103.3%
蝉鸣声	-0.5	-0.46	85.4%	97.5%
蝉鸣声	-0.5	-0.44	85.4%	96.8%
蝉鸣声	-0.5	-0.42	85.4%	94.3%

从此组实验可以看出,当对 g 值进行微调后,蝉鸣声的能量并没有呈现出巨大的变化。同时由于临界倾斜滤波器非线性补偿的特性,带来听觉上的感受不仅仅是增大或者减小音量这么简单,从而丰富了蝉鸣声的多样性。

3.2 对其它耳鸣声源的仿真

在常见耳鸣声源当中,除了蝉鸣声之外,还有水流声以及蚊子声等。那么这种方法是否也适用与这些声源呢,接下来用相同的实验方式来进行处理。其中 $\beta=0.35$,采样点数仍取10000个点,实验结果如下图所示:

1、水声

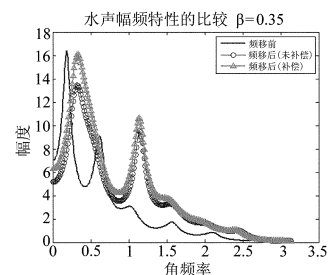


图10 幅频特性的比较

Fig. 10 Compare of amplitude-frequency characteristic

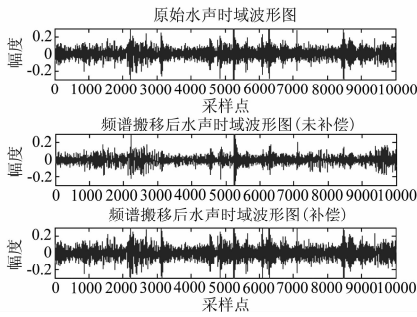


图 11 时域波形图的比较

Fig. 11 Compare of time domain waveform

2、蚊子声

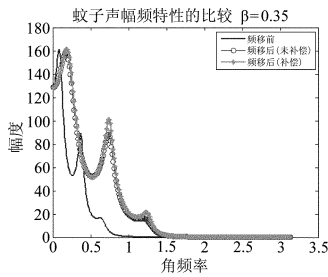


图 12 幅频特性的比较

Fig. 12 Compare of amplitude-frequency characteristic

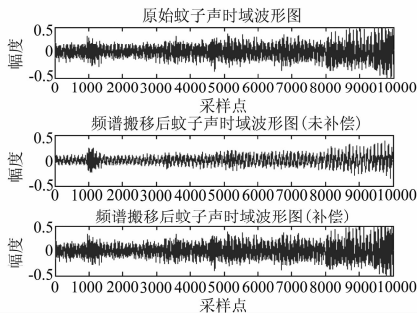


图 13 时域波形图的比较

Fig. 13 Compare of time domain waveform

从以上实验可以看出,双线性变换函数仍然适用于其它耳鸣声源,并可以对耳鸣声源的幅频特性进行改变。同时利用临界倾斜滤波器也能较好地调整声源能量,避免双线性变换函数不能合理控制共振峰幅度的缺陷,从而最终实现耳鸣声源的多样性。

由于篇幅有限,表 2 为常见耳鸣声源补偿前后能量百分比的统计:

表 2 不同耳鸣声源能量百分比的统计

Tab. 2 Statics of energy percentage with different tinnitus sound source

音频源	弯折系数 β	最佳系数 g 值	补偿前 $Per(g)$	补偿后 $Per(\hat{g})$
蝉鸣声	0.25	-0.5	85.4%	98.1%
蚊子声	0.25	-0.58	78.0%	96.3%
水声	0.25	-0.12	58.0%	97.6%
火声	0.25	-0.76	43.3%	95.7%
发动机	0.25	-0.38	55.3%	98.6%
螺旋桨	0.25	-0.18	61.2%	97.5%

从此组实验可以看出,针对其它耳鸣声源,临界倾斜滤波器的补偿效果还是比较好的,能够对损失的能量进行补偿。这也从侧面反映出该方法对其它声源也是适用的。

3.3 不同 β 值对耳鸣声源的影响

由于利用双线性变换函数来实现频谱搬移可能会导致幅频特性的恶化,从而对耳鸣声源带来音色畸变,特别是当 $|\beta|$ 值过大时这种效应会更加明显。那么接下来通过实验的方式来找出适用于耳鸣声源的 $|\beta|$ 取值范围。实验方法通过 10 人(5 男,5 女)对频谱搬移后(未补偿)的耳鸣声源进行测听来找出 $|\beta|$ 的极限取值。采样点数仍取 10000 个点。

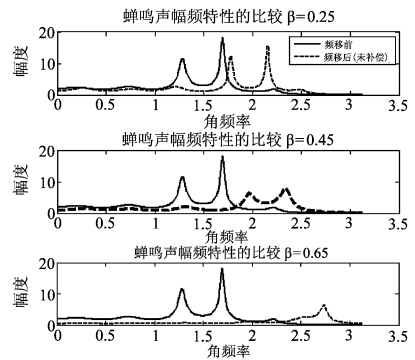


图 14 幅频特性的比较 $\beta=0.25、0.45、0.65$

Fig. 14 Compare of amplitude-frequency characteristic $\beta=0.25、0.45、0.65$

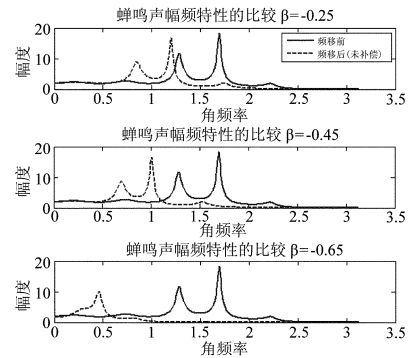


图 15 幅频特性的比较 $\beta=-0.25、-0.45、-0.65$

Fig. 15 Compare of amplitude-frequency characteristic $\beta=-0.25、-0.45、-0.65$

从图 14, 图 15 $|\beta|$ 值的取值趋势可以看出,当 $|\beta|=0.65$ 时,共振峰的特性已经严重恶化。而从 10 名特定实验对象的测听效果来说,7 人感受当 $|\beta|=0.58$ 左右时蝉鸣声产生失真,3 人感受当 $|\beta|=0.61$ 左右时蝉鸣声产生失真。这同实验的幅频特性上来看也是较为一致的。

利用同样的方法,对其它常见耳鸣声源进行实验后可以得出如表 3 的实验结果:

从表 3 可以看出,针对不同的耳鸣声源,其 β 的最大范围是不同的。但表中的数据仅为参考值,根据不同人的听觉感受可能存在一定的浮动。如果考虑到软

件设计的通用性,建议在利用双线性变换函数对耳鸣声源进行处理的时候, $|\beta|$ 的取值范围应控制在 0.6 以内,以避免幅频特性的恶化所带来的不良后果。

表 3 不同耳鸣声源的极限 β 值

Tab.3 The critical β value of different tinnitus sound source

音频源	极限弯折系数 $\beta \rightarrow 1$	极限弯折系数 $\beta \rightarrow -1$	补偿系数 g 值 (不补偿)
蝉鸣声	0.58	-0.60	0
蚊子声	0.87	-0.73	0
水声	0.65	-0.48	0
火声	0.87	-0.81	0
发动机	0.75	-0.67	0
螺旋桨	0.78	-0.74	0

4 结论

双线性变换函数是语音转换中一种重要的转换方法,本文借用这种思想通过 $|\beta|$ 值的动态调整来实现了耳鸣匹配声源的频谱搬移。同时也指出了利用传统双线性变换函数进行频谱搬移会带来能量衰减的不足,并采取临界倾斜滤波器的方案来对存在的问题进行弥补。

同时文中也提出了利用补偿前后能量的百分比来寻找最佳补偿系数 g 值的方法,并给出了 g 值微调的建议范围。总体来说,实验取得了良好的效果,丰富了耳鸣匹配声源的多样性,具有一定的实用参考价值。

参考文献

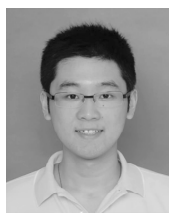
- [1] 王洪田. 耳鸣的诊断治疗新进展[J]. 实用医学杂志, 2005, 21(2): 114-116.
Wang H T. The Progress of Tinnitus Diagnosis and Treatment[J]. The Journal of Practical Medicine, 2005, 21(2): 114-116. (in Chinese)
- [2] 蒋涛. 耳鸣诊断和治疗技术新进展[J]. 临床耳鼻咽喉头颈外科杂志, 2009, 23(22): 1009-1010.
Jiang T. The Progress of Tinnitus Diagnosis and Treatment [J]. Journal of Clinical Otorhinolaryngology Head and Neck Surgery, 2009, 23(22): 1009-1010. (in Chinese)
- [3] 薛英,叶林峰,宋鹏. 听尼特耳鸣综合诊断治疗仪的临床应用[J]. 听力学及言语疾病杂志, 2007, 15(6): 490-495.
Xue Y, Ye L F, Song P. A Clinical Study of the Effectiveness of TinniTest in the Diagnosis and Management of Tinnitus [J]. Journal of Audiology and Speech Pathology, 2007, 15(6): 490-495. (in Chinese)
- [4] 赵力. 语音信号处理[M]. 北京:机械工业出版社, 2007, 56-57.
- [5] 孙燕,姜占才,王得芳. 语音频谱分析与应用[J]. 计算

机与现代化, 2010, 176(4): 200-202.

Sun Y, Jiang Z C, Wang D F. Analysis and Application of Voice Spectrum[J]. Computer and Modernization, 2010, 176(4): 200-202. (in Chinese)

- [6] 李波. 语音转换的关键技术研究[D]. 武汉:国防科学技术大学研究生院. 2005.
Li B. Studies on Key Techniques for Voice Conversion [D]. WuHan: Graduate School of National University of Defense Technology. 2005. (in Chinese)
- [7] H. Mizuno, M. Abe. Voice Conversion Algorithm Based on Piecewise Linear Conversion Rules of Formant Frequency and Spectrum Tilt [J]. IEEE International Conference, 1994, vol. 1: 469 -472.
- [8] 王京辉. 语音信号处理技术研究[D]. 山东: 山东大学研究生院. 2008.
Wang J H. Studies on Speech Signal Processing Techniques [D]. ShanDong: Graduate School of ShanDong University. 2008. (in Chinese)
- [9] 李凤莲,张雪英. 基于伪倒谱感觉加权滤波器的设计与应用[J]. 计算机工程, 2010, 36(3): 275-279.
Li F L, Zhang X Y. Design and Application of Perceptual Weighting Filter Based on Pseudo-cepstrum [J]. Computer Engineering, 2010, 36(3): 275-279. (in Chinese)
- [10] Cao Y Y, Zhao D Y, Zhou L J. Research on Order-Variable Code Exited Linear Prediction Speech Coding Method [J]. IEEE International Symposium, 2009, pages: 1-4.

作者简介



熊培杰 (1986-),男,四川大学电子信息学院硕士研究生。主要研究方向为语音/声学信号处理等。
E-mail: xpjyjt@gmail.com



何培宇 (1963-),女,四川大学电子信息学院教授,博士生导师。主要研究方向为语音/声学信号处理、自适应阵列信号处理及 DSP 技术等。
E-mail: hpysbsy@163.com



潘帆 (1982-),男,四川大学电子信息学院博士研究生。主要研究方向为语音/声学信号处理等。
E-mail: popfan229@gmail.com