

基于 ERB 尺度的心理声学模型及其数值计算

张伟豪^{1,2}, 许枫²

(1. 中国科学院声学研究所, 北京 100190; 2. 中国科学院研究生院, 北京 100036)

摘要: 粗糙度是评价声音幅度变化的一个基本的心理声学参数。针对人耳基底膜的滤波特性, 提出了一种基于等效矩形带宽(ERB)的心理声学粗糙度模型。模型通过划分 ERB 尺度为 75 个滤波通道, 计算幅度调制信号的广义调制指数, 然后通过加权滤波并求和得到总的粗糙度。数值计算结果验证了模型对于幅度调制信号粗糙度计算的准确性, 并且对于诸如调频信号、未调制信号等非幅度调制信号, 模型的计算结果也有较好的一致性。特别是对于非调制信号, 模型的计算结果与主观低粗糙感对应良好。

关键词: 心理声学; 粗糙度; 等效矩形带宽(ERB); 广义调制指数

中图分类号: B845.2

文献标识码: A

文章编号: 1000-3630(2011)-02-0161-06

DOI 编码: 10.3969/j.issn1000-3630.2011.02.011

Psychoacoustical roughness model and calculation method based on ERB-scale

ZHANG Wei-hao^{1,2}, XU Feng²

(1. Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China;

2. Graduate University of Chinese Academy of Sciences, Beijing 100036, China)

Abstract: As a basic psychoacoustic parameter, roughness is used to evaluate amplitude variations. According to filter characters of basilar membrane, roughness model is proposed based on equivalent rectangular bandwidth (ERB). By setting 75 filter channels on the ERB-scale, generalized modulation depth is calculated, weighted, filtered and summed to get the overall roughness. Numerical calculation results verify the validity of the model for calculating amplitude modulation signal's roughness. The proposed model also behaves better coherence to jury test data while being used for non amplitude modulation signals such as frequency modulation signal and unmodulated signals. Especially for unmodulated signals, the model ensures a correct calculated roughness in accordance with the subjectively sensed low roughness.

Key words: psychoacoustics; roughness; equivalent rectangular bandwidth (ERB); generalized modulation depth

0 引言

随着社会的发展, 各行各业对其周围声环境质量的要求越来越高, 不仅对各类产品的噪声限值提出了要求, 而且要求声品质(sound quality)^[1]要好。另一方面, 非致命性声武器逐渐成为人们关注的焦点。这种软杀伤武器利用某些特殊声音信号给人带来的烦恼感、粗糙感等心理声学效应以达到攻击的目的, 使敌方人员出现非致命机体及精神损伤, 可避免造成大量无辜人员的伤亡。

心理声学粗糙度可作为非致命性声武器攻击性能的客观评价参数之一, 它主要是用来描述声音信号调制程度的心理声学参数, 反映信号调制幅度的大小、调制频率的分布情况等特征。粗糙度的单

位是 asper, 定义声压级为 60 dB、频率为 1 kHz 的纯音信号, 在调制频率为 70 Hz、100%调制时的粗糙度为 1 asper。

目前, 粗糙度的计算尚无统一的国际标准。Von Helmholtz^[2]是第一个详细描述粗糙度这种听觉现象的学者。Aures^[3], Zwicker and Fastl^[4]和 Sottek^[5]均提出了各自的计算模型, 这些模型的基本差异在于临界频带上信号的包络变化怎样转化成特性粗糙度。

Aures 的粗糙度模型首先利用各个临界频带通道的激励包络计算广义调制深度(generalised modulation depth), 再将其转化成特性粗糙度, 然后在整个临界频带内求和得到总粗糙度。这种模型是最复杂也是最准确的, 它依赖于声音信号的调制指数、幅度调制信号的调制频率和载波频率等。对于非调制声音信号, 模型的计算结果与主观粗糙度测量结果也有较好的一致性。

Zwicker and Fastl 对 Aures 的粗糙度模型进行了改进, 利用调制频率和掩蔽深度(masking depth)

收稿日期: 2010-03-14; 修回日期: 2010-05-14

作者简介: 张伟豪(1984-), 男, 四川南充人, 博士研究生, 研究方向为心理声学、水声信号处理。

通讯作者: 张伟豪, E-mail: whz0717@sina.com

的乘积来计算粗糙度。Sottek 利用非线性压缩和自相关的激励包络来计算特性粗糙度,但是,他的模型计算结果和主观数据相比还存在较大的差异。

Aures 提出的粗糙度模型是在临界频带 Bark 尺度上完成的。后来,Moore 研究了耳蜗滤波器的频响特性,提出了等效矩形带宽(用 ERB 度量)的概念^[6]。ERB 尺度与 Bark 尺度在高频段差异不大,在低频段(500 Hz 以下)差别较大,ERB 尺度更精确地描述了人耳基底膜的频率选择特性。因此,本文的主要工作就是对 Aures 的粗糙度模型进行改进,在 ERB 尺度上计算粗糙度,并将计算结果与国外研究人员提供的实验数据对比,以验证模型的准确性。

1 粗糙度模型结构

在详细分析了 Aures 的粗糙度计算模型后,本文对 Aures 的模型进行了改进,主要表现在以下几个方面:一是改进了外耳、中耳的滤波效果;二是滤波通道是基于 ERB 尺度划分的;三是利用 ERB 听觉滤波器替代 Aures 模型中的非线性梯形滤波器(nonlinear trapezoid filter)。图 1 给出了改进后粗糙度模型的结构及仿真计算流程。

整个模型的计算可分为四个阶段实现。第一阶段是声音信号的预处理,包括信号时域加窗截取、FFT 变换、外耳中耳滤波以及忽略小于绝对听阈的频率成分;第二阶段是通道激励信号 $e_i(t)$ 的计算,包括 ERB 尺度的划分、ERB 听觉滤波器设计、各通道滤波,并计入相位信息,IFFT 变换得到 $e_i(t)$;第三阶段是计算广义调制指数 m_i^* ,包括包络谱的计算、与调制频率有关的滤波器滤波,并转换到时域计算 m_i^* ;第四阶段是特性粗糙度 r_i 的计算,通过平方并加权 m_i^* 实现,包括载波频率相关的加权、通道间互相关性加权。

最后,总粗糙度 R 即是所有通道上计算的特性粗糙度值的和。

2 粗糙度模型数值计算流程

2.1 声音信号的预处理

模型输入端为采集的时域声音信号,首先对其作以下几方面的预处理:

(1) 读取 200 ms 的声信号,利用 Blackman 窗函数加窗,作 FFT 变换,将时域信号转换成频域信号。

(2) 外耳、中耳滤波。将频域信号的各个频率成分通过外耳、中耳传递函数进行滤波。图 2~图 3

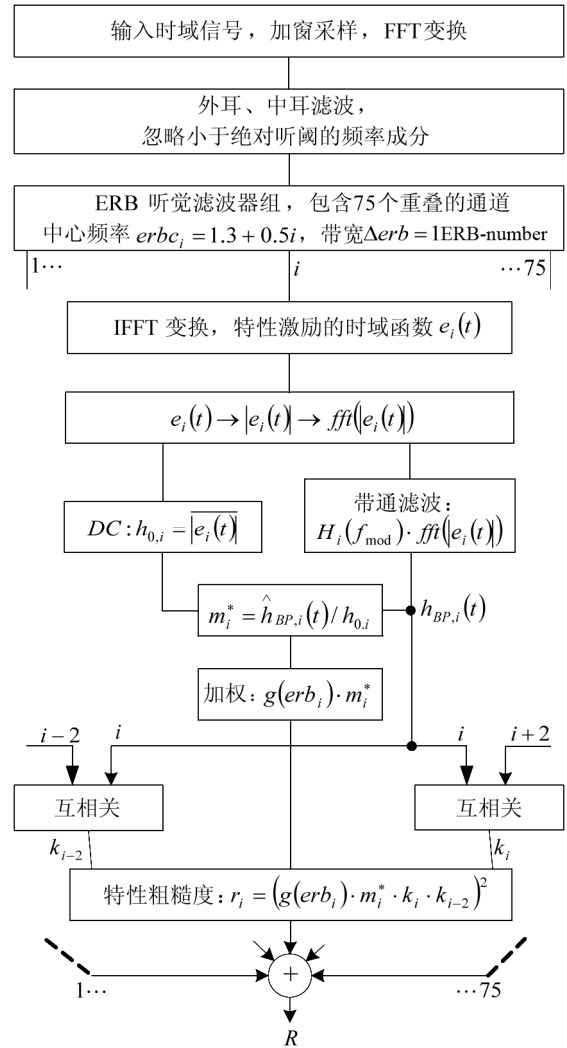


图 1 粗糙度模型结构

Fig.1 Structure of roughness model

分别是通过内插美国标准 ANSI S3.4-2005^[7]中的相关数据得到的外耳自由场、混响场传递函数和中耳传递函数。声音在传递过程中,通过外耳后声压级增加了;通过中耳后声压级减小了。图 2 中的曲线表示声音信号通过外耳后的声压级减去声音信号在自由场(混响场)中的声压级;图 3 中的曲线表示声音信号通过外耳后的声压级减去通过中耳后的

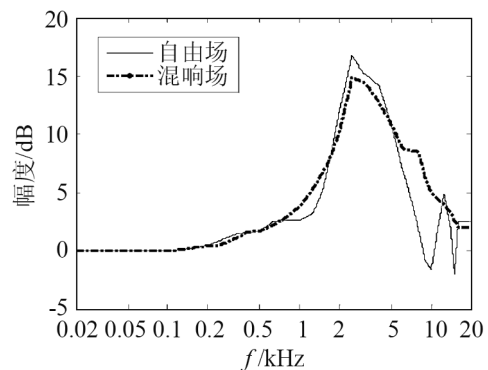


图 2 外耳传递函数

Fig.2 Transfer function of outer ear

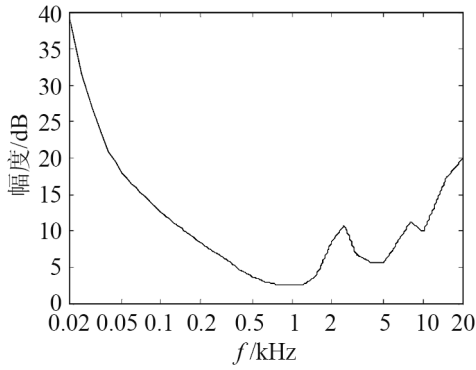


图3 中耳传递函数
Fig.3 Transfer function of middle ear

声压级。

(3) 忽略小于绝对听阈的频谱级。绝对听阈 L_{iq} 指人耳能听到声音的最小分贝级, 对小于绝对听阈的频率成分, 人耳是感觉不到粗糙度的。因此, 比较每个频率成分处的谱级与这个频率处的绝对听阈, 只保留大于或等于绝对听阈的谱级参与后续计算。

2.2 激励信号的计算

根据 Moore 的论述, 人耳听觉滤波器的中心频率只覆盖 50~15000 Hz(对应 ERB 尺度范围为 1.8~38.8, 转换关系参考文献[7])的频率范围。因此, 理论上我们需要设计这一范围内每一个频率处的听觉滤波器形状, 但实际上是没有必要的, 只需以 ERB 为间隔确定每一 ERB 带内中心频率处的滤波器即可。

ERB 听觉滤波器设计包括计算每个频率成分的 ERB 级、分别确定中心频率上下的滤波器形状函数, 具体过程这里不再详述, 读者可参考文献[7-8]。图4给出了所有频率处的 ERB 级^[7]为 100dB 的听觉滤波器组 MATLAB 设计结果, 这组滤波器的中心频率 $erbc$ 为 $1.8+(i-1)ERB\text{-number}$, $i=1, 2, \dots, 38$, 共计 38 个滤波器, erb 表示 ERB 尺度变量, ERB-number 表示变量 erb 的单位。

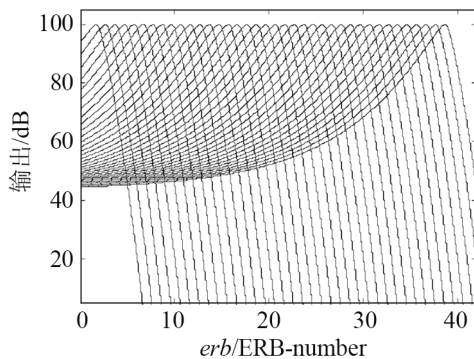


图4 ERB 听觉滤波器组
Fig.4 ERB Aural filters bank

本文提出的粗糙度模型设计了 75 个重叠的带宽为 1 ERB 的滤波通道, 每个通道在 ERB 尺度上的中心频率 $erbc_i$ 为:

$$erbc_i = 1.3 + 0.5i, \quad i=1, 2, \dots, 75 \quad (1)$$

然后利用听觉滤波器在 $[erbc_i - 0.5, erbc_i + 0.5]$ 上滤波, 得到每个通道的特征激励谱。

滤波器的设计以及对信号进行滤波均是在频域内完成的, 因此需要对滤波器的输出进行逆傅立叶变换(IFT)得到 200ms 长的激励信号 $e_i(t)$, 作 IFT 时需要计入原始数据帧的相位信息。

2.3 广义调制指数的计算

广义调制深度 m_i^* 的值可通过如下方式计算获得: 首先对激励信号 $e_i(t)$ 的包络谱加权滤波; 其次将滤波后的包络谱变换到时域, 计算时域信号的 RMS 值; 然后将这个 RMS 值比上 $e_i(t)$ 包络信号的直流值得 m_i^* 。

激励信号包络的变化包含在包络谱 $DFT(|e_i(t)|)$ 的低频部分中, 并且粗糙度将随调制频率的变化表现出带通特性, 为了建模粗糙度的这种特性, 利用调制频率相关的加权函数 $H_i(f_{mod})$ ^[3]对激励信号的包络谱进行滤波处理。

图5给出了当频率为 200Hz、920Hz、7700Hz(对应 erb 分别为 5.8、15.0、32.9)时的加权函数随调制频率的变化曲线^[3]。其他频率处的加权函数由如下方法确定:

- (1) 当 $erb < 5.83$ 时, $H_i(f_{mod})$ 等于 200Hz 处的加权函数;
- (2) 当 $5.83 < erb < 17.44$ 时, $H_i(f_{mod})$ 取 920Hz 处的加权函数值;
- (3) 当 $17.44 < erb \leq 38.9$ 时, $H_i(f_{mod})$ 取 7700Hz 处的加权函数值。

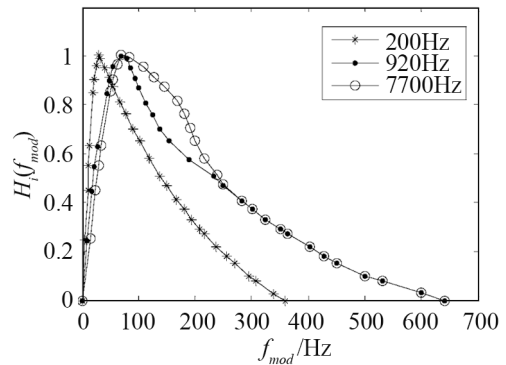


图5 调制频率相关的加权函数
Fig.5 Weighting function reference to modulation frequency

确定了每个通道的加权滤波函数后, 对激励信号的包络滤波, 然后通过逆傅立叶变换得到滤波后的包络信号 $h_{BP,i}(t)$ 。广义调制深度可由式(2)确定:

$$m_i^* = \begin{cases} \frac{\tilde{h}_{BP,i}(t)}{h_{0,i}} & \text{当 } \frac{\tilde{h}_{BP,i}(t)}{h_{0,i}} \leq 1 \text{ 时} \\ 1 & \text{当 } \frac{\tilde{h}_{BP,i}(t)}{h_{0,i}} > 1 \text{ 时} \end{cases} \quad (2)$$

其中, $\tilde{h}_{BP,i}(t)$ 表示包络信号 $h_{BP,i}(t)$ 的均方根 RMS, $h_{0,i} = \overline{|e_i(t)|}$ (时间平均)。

2.4 特性粗糙度的计算

载波频率对粗糙度的影响主要是通过引入加权函数 $g(z_i)$ [3] 实现的, $g(z_i)$ 取值区间为 0.5~1.2。文献[3]中的 $g(z_i)$ 是在 Bark 尺度上定义的, 这里需要将其转换到 ERB 尺度上(转换的桥梁是频率), 转换后的加权函数 $g(erb_i)$ 如图 6 所示。

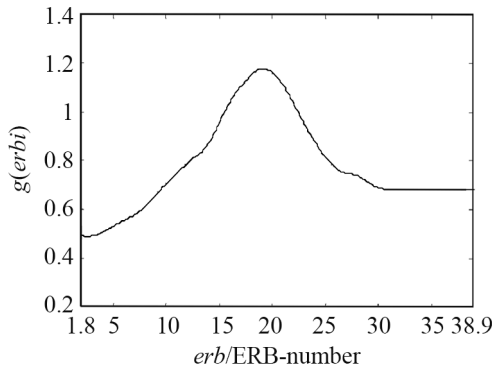


图6 载波频率相关的加权函数
Fig.6 Weighting function reference to carrier frequency

如果通道之间的调制包络信号不相关, 简单的求和 $(g(erb_i) \cdot m_i^*)^2$ 将过高地估计总粗糙度, 如: 未调制信号、白噪声信号等的粗糙度接近于零, 即这种信号给人没有粗糙感, 但按以上方法计算出的总粗糙度将不为零。因此需计入通道间包络信号互相关性的影响, 计算通道 $i-2$ 与 i 的互相关系数 k_{i-2} 以及通道 i 与 $i+2$ 的互相关系数 k_i (注: 当 $erbc=1.8, 2.3$ 以及 $erbc=38.3, 38.8$ 时, 在公式(2)中分别只计算 k_i 以及 k_{i-2} 。), 则特性粗糙度可由式(3)表示:

$$r_i = (g(erb_i) \cdot m_i^* \cdot k_{i-2} \cdot k_i)^2 \quad (3)$$

则总粗糙度 R (单位: asper) 可表示为:

$$R = C \cdot \sum_{i=1}^{75} r_i \quad (4)$$

其中的校准因子设置为 $C=0.277$, 目的是为了使得声压级为 60 dB、频率为 1 kHz 的纯音信号, 在调制频率为 70 Hz, 100% 调制时的粗糙度为 1 asper。

3 计算结果与比较

3.1 幅度调制纯音

Aures 通过实验以及建模分析表明, 随着调制

频率 f_{mod} 的变化, 幅度调制(AM)纯音信号的粗糙度表现出带通特性^[9]。下面利用本文提出的模型计算 AM 信号的粗糙度, 并与 Zwicker 在其著作^[4]中给出的数据进行对比。计算用 AM 信号: 声压级 60 dB、100% 幅度调制纯音信号。图 7 分别绘出了载波频率 $f_c=0.25, 1, 4, 8$ kHz 时, 粗糙度随调制频率的变换曲线。从图中可看出, 模型计算数据与 Zwicker 的数据有较好的一致性, 并且表现出了带通特性。

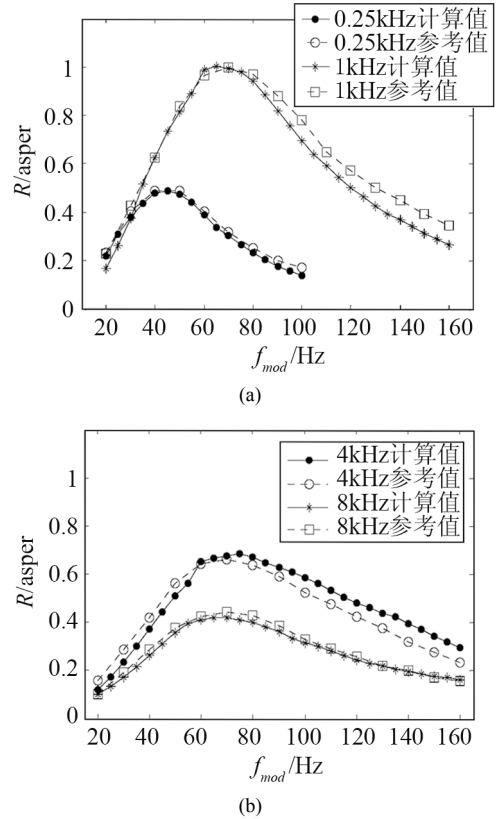


图7 AM 信号粗糙度计算结果
Fig.7 Calculated roughness results of AM signals

3.2 频率调制纯音

本文的粗糙度模型虽然是针对幅度调制(AM)信号提出的, 但是对于频率调制(FM)信号的计算也有一定的准确性。Kemp^[10]研究了 FM 信号的粗糙度特性, 并提供了大量的主观实验数据, 可作为检测模型仿真计算结果准确性的参考数据。调制频率、频偏以及声压级是影响 FM 信号粗糙度的三个重要因素, 下面分别加以讨论。

图 8 分别给出了 FM 信号的相对粗糙度随调制频率 f_{mod} 、频偏 Δf 以及声压级 SPL 的变化趋势, 以及与参考实验数据^[10]的对比。其中, R_0 为声压级 60 dB、频率 1.6 kHz 的纯音在调制频率 70 Hz、频偏 800 Hz 调制时的粗糙度。

分析图 8(a)可看出, 除了调制频率在 20~60 Hz 范围内, 模型计算粗糙度与 Kemp 的实验数据存在

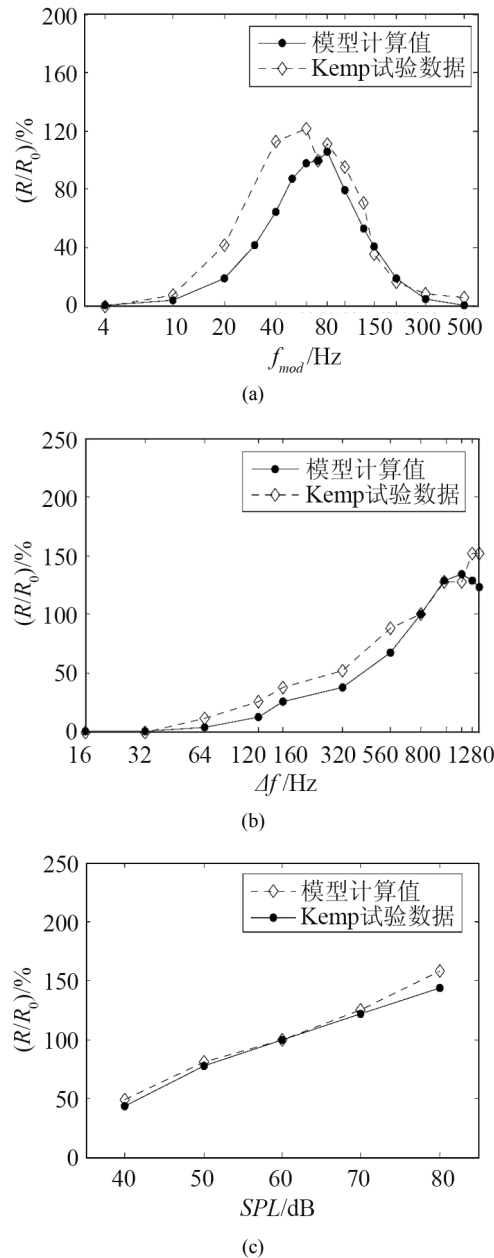


图8 FM信号相对粗糙度计算结果
Fig.8 Calculated relative roughness results of FM signals

较小的差异外,其他调制频率处的粗糙度与实验数据的一致性均较好;并且带通特性良好。图8(b)、8(c)的结果也显示了模型的计算值与实验数据一致性较好,极个别的点除外(例如:图8(c)中声压级为80 dB时,两者有较小差异)。

3.3 调制指数的影响

仿真信号:调制频率为70 Hz、声压级为70 dB,载波频率为1 kHz的AM信号,调制指数范围 $0 \leq m \leq 1$,变化步长0.05。图9是模型计算的粗糙度随调制指数 m 的变化趋势,并且与理论计算值 $R=1.36m^{1.6}$ (Zwicker认为幂指数 $p=1.6$ 与实验数据近似最好)^[4]作对比。从图中可看出,模型的计算结

果与理论值有较好的一致性。

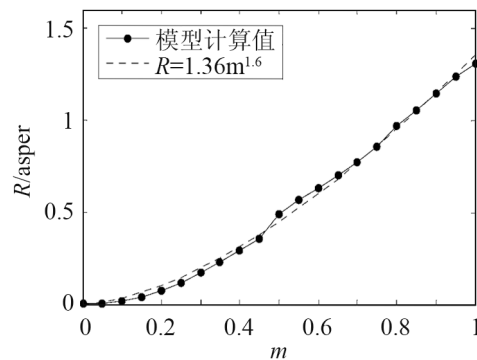


图9 调制指数的影响
Fig.9 Effect of modulation index

3.4 未调制信号

由于工业噪声和自然界的声信号难免包含各式各样的未调制信号,因此,未调制信号粗糙度的计算是检测模型正确性的有效途径。文献[4]指出,能被人耳感知的粗糙度的门限是 $R_h=0.07$ asper,即是说,对于未调制信号来说,模型计算的粗糙度应该小于门限值 R_h 。表1给出了几种不同形式的未调制信号的粗糙度计算结果,由此可看出,本文提出的粗糙度模型也适用于未调制信号的计算。

表1 未调制信号粗糙度				
Table 1 Roughness of unmodulated signals				
信号类型	单频(2 kHz)	单频(3.4 kHz)	白噪声	粉红噪声
粗糙度/asper	0.0005	0.0006	0.0167	0.0482

4 结论

粗糙度的计算目前国际上还没有统一的标准,许多学者对粗糙度进行了主观实验研究以及建模计算,并得到一系列结果。本文在总结前人工作的基础上,详细给出并论述了在ERB尺度上的粗糙度模型结构及计算流程。在利用提出的模型计算粗糙度的过程中,对于不同的声信号种类(AM、FM、噪声等),参考了不同学者的主观实验数据。对比发现,除个别情况外(例如:图7(b)、图8(a)中的几个点),本文的仿真计算数据与参考数据具有较好的一致性。

本文的模型是专门针对幅度调制声信号提出的,但是模型计算以及与实验数据的对比表明,该模型同样适用于其他声信号粗糙度的计算,并且计算结果的准确性较好。同时可以看出,模型的计算流程是比较复杂的,计算量较大,下一步工作将是对模型的某些环节进行优化,减少计算量,以满足实际中快速分析声信号心理声学参数特征的需要。

参 考 文 献

- [1] Tsuge K, Kanamaru K. A study of noise in vehicle passenger compartment during acceleration[R]. Proceeding of the 1985 SAE Noise and Vibration Conference, SAE850965.
- [2] Helmholtz H. Die lehre von den tonempfindungen(The principle of tone perception)[D]. Wissenschaftliche Buchgesellschaft, Darmstadt, unveränderte Auflage, 1968, 7.
- [3] Aures W. Ein Berechnungsverfahren der Rauigkeit(A Procedure for Calculating Auditory Roughness)[J]. Acustica, 1985, **58**(5): 268-281.
- [4] Zwicker E, Fastl H. Psychoacoustics: Facts and Models[M]. Third edition, Germany: Spring-Verlag Berlin Heidelberg, 2007.
- [5] Sottek R. Modelle zur Signalverarbeitung im menschlichen Gehör(Models of signal processing in human hearing)[D]. Dissertation, RWTH Aachen, Germany, 1993.
- [6] Moore B C J, Glasberg B R, Baer T. A model for the prediction of thresholds, loudness and partial loudness [J]. Journal Audio Eng. Soc., 1997, **45**(4): 224-237.
- [7] ANSI S3.4-2005, Procedure for the Computation of Loudness of Steady Sounds[S].
- [8] 马元峰, 陈克安, 王娜. Moore 响度模型的数值计算方法[J]. 声学技术, 2008, **27**(3): 390-395.
MA Yanfeng, CHEN Ke'an, WANG Na. The study of loudness' calculation based on moore's model[J]. Technical Acoustics, 2008, **27**(3): 390-395.
- [9] Aures W. Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, ein Beitrag zur gehorbezogenen Schallanalyse[D]. PhD thesis, TU München, 1984.
- [10] Kemp S. Roughness of frequency-modulated tones[J]. Acustica – Acta Acustica, 1982, **50**(2): 113-123.