

真实听觉空间研究中 FETF 的 测量和估计方法^①

吴 镇 扬

(东南大学无线电系 南京, 210018)

陈 家 澍

(威斯康星大学神经生理系)

理查德 A. 雷阿尔

(威斯康星大学神经生理系)

1994 年 6 月 10 日收到

摘要 自由声场至耳膜的传递函数 (FETF) 被认为是影响辨别声源方向的主要因素之一。FETF 已被用于真实听觉空间的计算机仿真。对于 FETF 的估计和测量方法, 本文提出了具体的改进意见。对于单脉冲激励的测量方法, 我们建议采用 LMS 估计方法, 计算机模拟结果表明 LMS 方法优于经典的 FFT 方法。文章中还对各种用于测量 FETF 冲激响应的激励方法作了分析和比较。

Measurement and estimation of free field to eardrum transfer function for virtual acoustic space research

WU Zhenyang

(Department of Radio Engineering of Southeast University Nanjing, 210008)

CHEN Jiashu

(Department of Neurophysiology of University of Wisconsin-Madison)

Richard A. Reale

(Department of Neurophysiology of University of Wisconsin-Madison)

Received June 10, 1994

Abstract Free field to Eardrum Transfer Function (FETF) is one of the major factors influencing the identification of the sound source direction. The FETF can be employed to generate a Virtual Acoustic Space (VAS) by computer and other equipment. In this paper we have approached the methods to improve the measurement and estimation of FETF. LMS method was suggested to estimate FETF for single pulse measurement. The computer simulation results have shown that LMS method is much better than empirical FFT method. This paper also gives a simple description of excitation signals for measuring the impulse response of FETF.

一、前 言

人类以及大部分动物均具有不同程度的空间听觉能力。从而, 他们可以确定声源的方位, 甚至在噪声环境中集中注意力于一特殊的目标。长期以来人们在声源定位的研究方面基本上只注意

^①获江苏省自然科学基金与国家教委留学回国人员基金资助

研究耳间时间差 (ITD) 和耳间声级差 (ILD)。早在三十年代, Stevens 和 Newman 就提出了声源定位的双工理论 (Duplex Theory)^[1]。这一理论指出 ITD 和 ILD 是影响空间声源定位的主要因素, 这是两耳分开的一个自然结果。在确定声音的方位方面, 双工理论依然不能解释许多听觉现象。这包括在单耳条件下的听觉定位机理、声源在两耳对称平面上的听觉定位机理以及在对称平面上前后分辨的机理。事实上由于头、肩、耳廓等部位的互相作用, 抵达耳膜的人射声波已发生了变化。此变化的特性取决于入射的方向。Butler^[2]等也指出: 这样一个传输过程所引起的谱特性的变化在空间声源的定位中起着重要的作用。这一传递函数被定义为自由声场至耳膜的传递函数 (Free-field to Eardrum Transfer Function), 简称 FETF。在心理学研究中也定义为与头部关联的传递函数 (Head Related Transfer Function)^[3], 简称 HRTF。

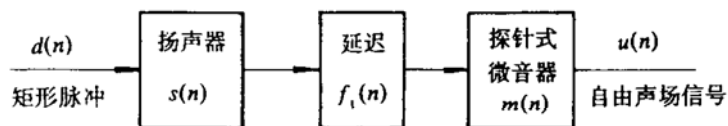
在得知 FETF 后, 加上 ITD 和 ILD, 我们便可以用计算机去实现这一真实听觉空间 (Virtual Acoustic Space), 简称 VAS。这一系统通常是用耳机或插入式耳机直接将声音送入左右耳。这一技术可用于多个领域, 例如, 建筑声学、听觉生理学、听觉心理学、计算机接口技术、计算机幻境技术以及用于减轻视觉负担的导航辅助技术等^[4-7]。在听觉生理学的研究中, 我们称这一系统为闭路声学系统, 用之代替自由声场的声激励信号^[4]。这一系统由主计算机、数/模和模/数转换器、音频放大器以及插入式的耳机组成。真实听觉空间的信号由声源与 FETF 卷积生成。这一系统的主要优点是它可独立地控制各种听觉参数, 例如, ITD、ILD 及频谱特性等, 以利于探讨听觉神经系统在声音定位中的机理。另外, 这一系统也大大地简化了实验室的声学条件。

由于 FETF 在空间听觉研究中的重要作用, 不少实验室着手在动物和人的耳膜处采集听觉数据。通常, 分别在空间各点发送同一宽频带的信号, 并同时用两只探针式微音器 (Probe Tube Microphone) 插入左右耳的耳道中采集听觉数据。这样一组听觉数据可用于估计不同方向的 FETF。然而, 由于测量噪声的影响以及不同方位采集听觉数据的动态范围相差较大 (约 20 dB)。如何从采集的数据中正确地估计 FETF, 以及如何进一步改进测量方法均是值得探讨的问题。

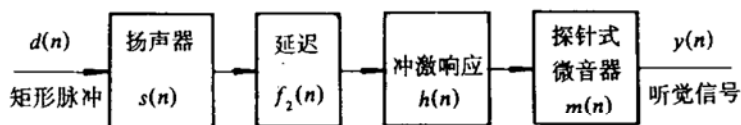
本文的第二节将具体地介绍 FETF 的测量方法。在已有的测量数据基础上, 我们对经典的 FETF 的估计方法和本文引用的 LMS 方法进行了比较, 并对不同信噪比下 FETF 的均方误差作了计算机模拟。根据这些结果, 在第三节中我们对今后的测量方法提出了改进意见。

二、FETF 的测量与估计

Musicant 及其同事^[8]在一无回声的实验室中应用矩形脉冲信号去激励空间位置可变的扬声器, 产生自由声场信号。扬声器的位置在一球面座标系统中变化, 其水平方向可以在 360° 范围内变化, 即从 -180° 至 $+180^\circ$; 垂直方向可以在 120° 范围内变化, 即从 -36° 至 $+90^\circ$ 。耳间的对称点置于球面的对称中心, 前方为 0° , 后方为 $\pm 180^\circ$ 。扬声器位置变化的步距为 4.5° 。与此同时在一麻醉后的动物的耳膜处采集听觉信号。其方法是用手术将探针式微音器植入动物耳道中接近耳膜的部位。在发送激励信号的同时采集左右耳的听觉信号, 这样得到了 3600 多个信号。图 1 描述了自由声场信号 $u(n)$ 与耳膜信号 $y(n)$ 的采集过程。由计算机产生的矩形脉冲 $d(n)$ 被加到自由声场扬声器, 得到输出 $s(n)$, 在没有动物的情形下, 自由声场激励信号 $u(n)$ 由 Knowles 探针式微音器测得。该微音器置于一假想目标的耳膜处, 其传递特性为 $m(n)$, 声延迟为 $f_1(n)$, 类似地在采集耳膜处的信号时, $y(n)$ 是由同一探针式微音器得到的, 不同的是微音器被植入麻醉后的动物耳道中, 并接近于耳膜处, 这里的声延迟为 $f_2(n)$, FETF 的冲激响应为 $h(n)$ 。在图 1 中, A/D 与 D/A 产生的影响被忽略。因此, 在 $u(n)$ 和 $y(n)$ 的基础上确定 $h(n)$ 是本节所要讨论的问题。由这些数据估计 FETF 是一系统辨识的问题, 即



(a) 自由声场激励信号的采集



(b) 耳膜处听觉信号的采集

图 1 FETF 的测量原理框图

$$y(n) = u(n) * h(n) \quad (1)$$

这里“*”表示作卷积运算, (1) 式可以用 FFT 求得

$$H(\omega) = \frac{Y(\omega)}{U(\omega)} \quad (2)$$

这里 $H(\omega)$ 、 $Y(\omega)$ 和 $U(\omega)$ 分别为 $h(n)$ 、 $y(n)$ 和 $u(n)$ 的离散付里叶变换。(2) 式通常定义为经典的传递函数估计方法。在这一经典方法中, 耳膜附近采集的信号就看作系统输出 $y(n)$ 。

然而, 由于扬声器的有限带宽特性, $u(n)$ 为一带限信号。在通带内, $u(n)$ 的特性也不平坦。图 2 为典型的自由声场激励信号, 受到矩形脉冲激励的扬声器输出却为衰减振荡波形, 见图 2(a)。其幅频特性见图 2(b), 在 2 kHz~30 kHz 范围内也有着 15 dB 的起伏。由于采用同一激励信号, 来自耳朵相反方向的声音抵达耳膜后将比较微弱, 在测量噪声的影响下 (主要来自微音器的热噪声), 采集的听觉信号的信噪比将比较差, 见图 3。所估计出的 FETF 实质上为

$$H(\omega) = \frac{Y(\omega)}{U(\omega)} = \frac{\bar{Y}(\omega) + \eta(\omega)}{U(\omega)} \quad (3)$$

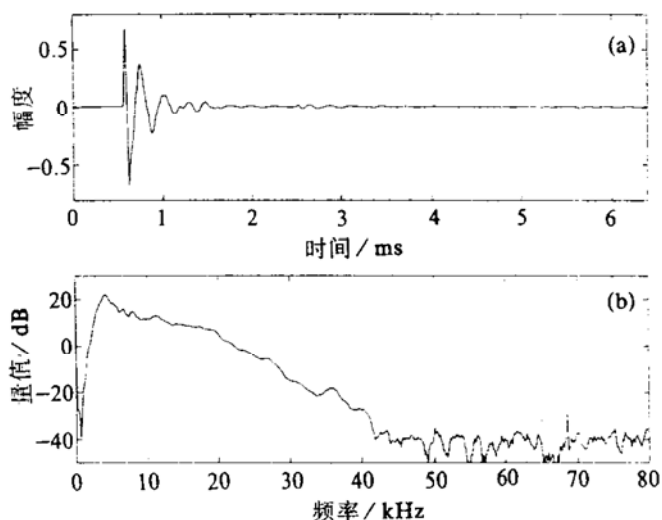


图 2 典型自由声场激励信号的
时域波形(a)与幅频特性(b)

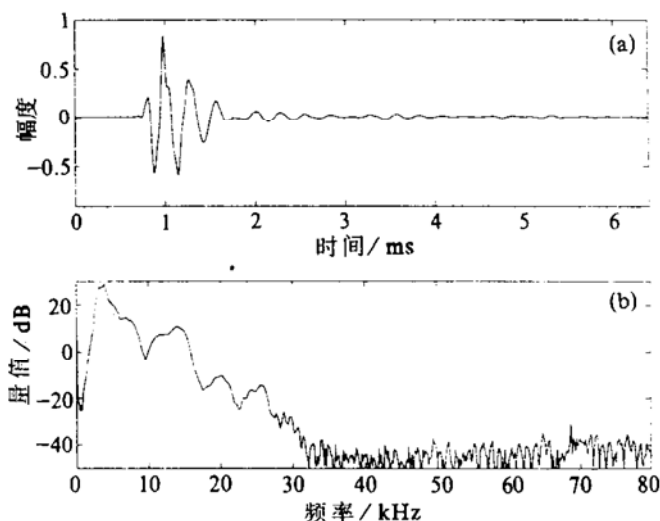


图 3 声源的水平方向为 -180° , 垂直方向
为 18° , 猫的左耳的听觉信号的时域
波形(a)与幅频特性(b)

式中 $\bar{Y}(\omega)$ 为真实的听觉信号, $\eta(\omega)$ 为测量噪音。图 4(a)为经典方法估计出的 FETF, 由图可见

在高于 30 kHz 和低于 2 kHz 的频段上, $\eta(\omega)$ 的份量起主导作用, 导致 $H(\omega)$ 在这一频段内的幅值高于我们所关心的 2 kHz~30 kHz 频段内的幅值。通常为了抵消测量噪声在这些频段上的影响, 采用数字带通滤波器滤除有效带宽以外的噪声。而由于各方向上 FETF 的频率特性是不一样的, 这一带通滤波器的上边界频率也是不确定的, 它在 30 kHz 附近变化。而在 2 kHz 至 30 kHz 的频率范围内, 测量噪声同样影响了 FETF 的真实特性, 尤其是在一些谷点频率附近。谷点频率是 FETF 的重要特征。

为了改进 FETF 的估计质量, 我们假设 FETF 是一有限冲激响应系统, 即

$$y(n) = h(n) * u(n) = \sum_{i=0}^{M-1} h(n-i) u(i) \quad (4)$$

式中 $h(n)$ 为 FETF 的冲激响应。从图 3(a) 可见, 听觉信号的主峰出现在 1 ms 处, 在 2 ms 处出现了第一个回声, 在 3.5 ms 处出现了第二个回声, 这些回声均是由于支撑动物的支架的声反射所引起的 (尽管这些支架均用吸声材料包装了起来, 还是有一定的声反射)。我们可以推断信号的有效宽度在 1.5 ms 左右, FETF 有限冲激响应的长度小于 1.5 ms。在 160 kHz 的采样速率下, 我们取(4)式中的 $M=256$ 。一般来说, 我们可按照最小平方滤波准则求得 $h(n)$ ^[9] 使得希望输出信号 $\bar{y}(n)$ 与实际采集的听觉信号 $y(n)$ 的误差能量 $Q = [\bar{y}(n) - y(n)]^2$ 最小。然而不幸的是这一运算必须估计 FETF 系统输入信号的自相关函数和输入信号与输出信号的互相关函数。对于图 2 所示的单脉冲激励信号, 用其自相关函数 (见图 5(a)) 构成的 256×256 自相关矩阵是一对称的 Toeplitz 矩阵。观察其特征值 λ_k , $k=1, 2, \dots, 256$, 其中 $\lambda_1 = 143$, $\lambda_2 = 143$, $\lambda_3 = 9.9$, ..., 而大部分特征值等于零或接近于零, 见图 5(b), 所以输入信号的自相关矩阵是一奇异矩阵。我们无论是用迭代的方法求解 Toeplitz 方程或对自相关矩阵直接求逆均会引入较大的误差。在大量的 FETF 计算中更难控制其误差。

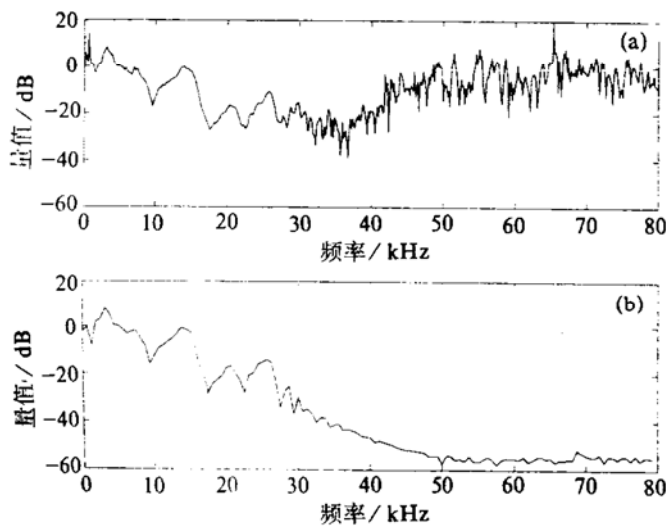


图4 FETF 的幅频特性。(a) 经典方法估计出的 FETF. 在低于 2 kHz 和高于 30 kHz 的频段上由于信号分量很弱, 测量噪声将起主要作用。(b) LMS 方法估计出的 FETF. 在输入信号微弱的频段上幅频特性呈自然下降。

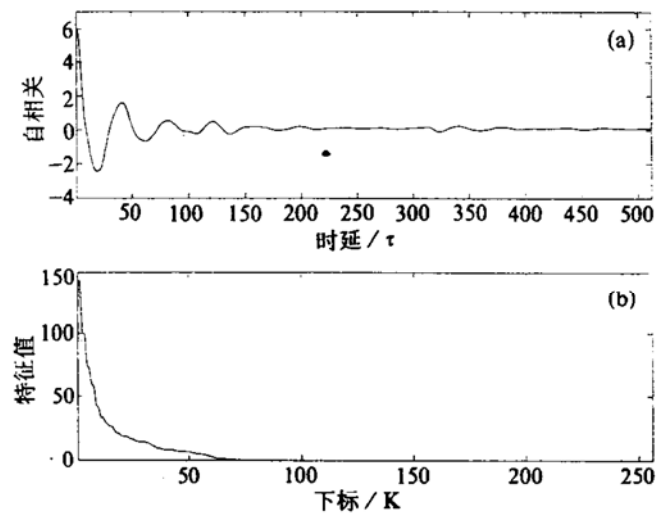


图5 自由声场激励信号的自相关函数(a)和自相关矩阵的特征值(b)。

为此, 我们按照最小均方准则求解 $h(n)$ 。为了保证有足够长的序列样本, 我们将 $u(n)$ 与 $y(n)$ 作周期性延拓。采用 Widrow^[10] 提出的最陡下降法求解 $h(n)$, 这一算法通常用于设计自适应

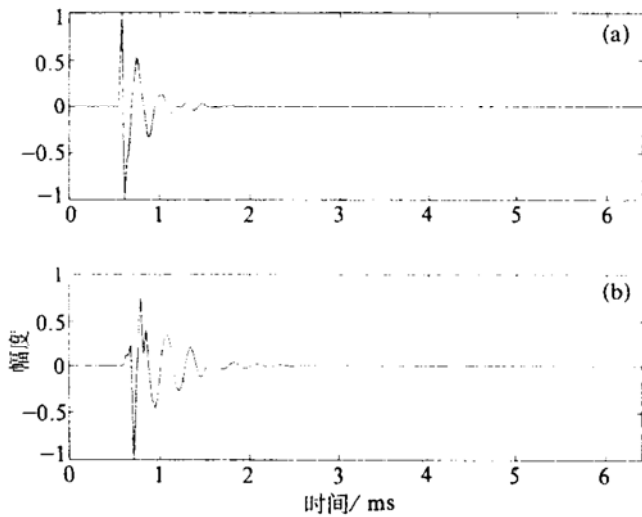


图6 用于计算机模拟的一组信号, (a) 自由声场激励信号; (b) 噪声和回波被滤除的听觉信号

应 FIR 滤波器, 它具有稳定、计算效率高、不需求解相关函数等特点。实践证明, 利用这一算法求得的 FETF 对噪声不敏感。我们可以看到, 在输入信号没有能量或能量很小的频段上, 其 FETF 的幅值是自然地变小的, 见图 4(b)。在我们所关心的频段内, 为了比较测量噪声对两种不同方法的影响, 我们用图 6 的输入输出信号做了计算机模拟, 即在输出信号 $y(n)$ 上迭加不同功率的带限噪声, 噪声的频带在 2 kHz~30 kHz 范围内均匀分布。在同一噪声功率下, 产生 20 组随机噪声分别迭加到输出信号上, 再用两种不同方法分别估计其 FETF 并计算此 FETF 与噪声功率为零时的 FETF 之间的均方误差。全部模拟工作在 VAX 的 Alpha 工作站上进行。图 7 给出了两种

不同方法在不同信噪比下估计出的 FETF 的均方误差的统计曲线。不难看出, 在信噪比较小的情形下, LMS 算法估计出的 FETF 其均方误差比经典方法要小得多。例如, 在信噪比为 30 dB 时, LMS 算法的归一化均方误差为 1%, 而经典方法的归一化均方误差大于 1.5%; 在信噪比等于 20 dB 时差别更大, LMS 算法的归一化均方误差小于 5%, 而经典方法的归一化均方误差已超过 25%。在信噪比大于 37 dB 时, 经典方法的均方误差将逐渐减少并趋于零, LMS 算法的归一化均方误差趋于 0.3% 的常数, 这是由 LMS 算法的失调量引起的。

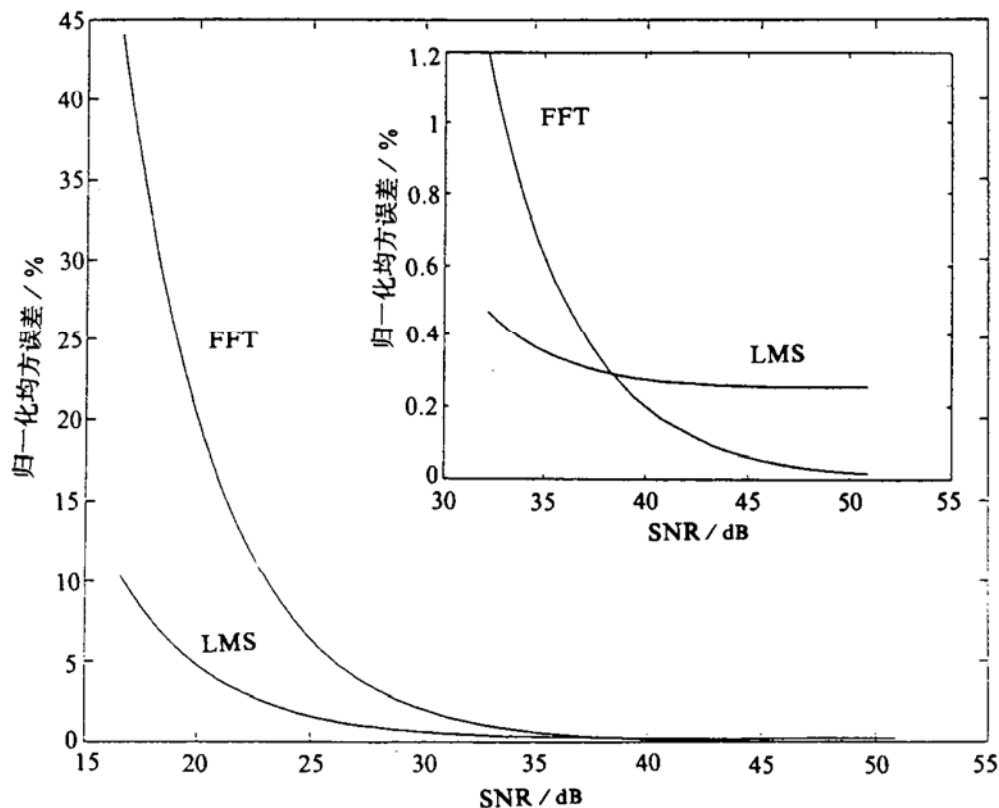


图7 在不同信噪比下, 两种不同方法估计出的均方误差的统计曲线。

图 8 给出了自由声场激励信号在空间三个方位上时实际记录的右耳听觉信号 (左边一列) 和在相同方向上用 LMS 算法估计的 FETF 所恢复的信号 (右边一列)。恢复信号是由原始的自由声场激励信号与各方位上 FETF 分别卷积后得到的。图中, AZ 表示水平方位, EL 表示垂直方

位。图中各方向上恢复信号与实际记录信号的相关系数均大于 0.99。对于 1816 个方向上的左右耳听觉信号, 我们全部用 LMS 算法计算其 FETF 的冲激响应, 并用这些冲激响应与自由声场激励信号一一卷积, 将其结果与原始记录信号逐一比较。所有恢复信号与原始记录信号的相关系数均不小于 0.98; 左耳的相关系数平均值为 0.9922; 右耳的相关系数平均值为 0.9931。

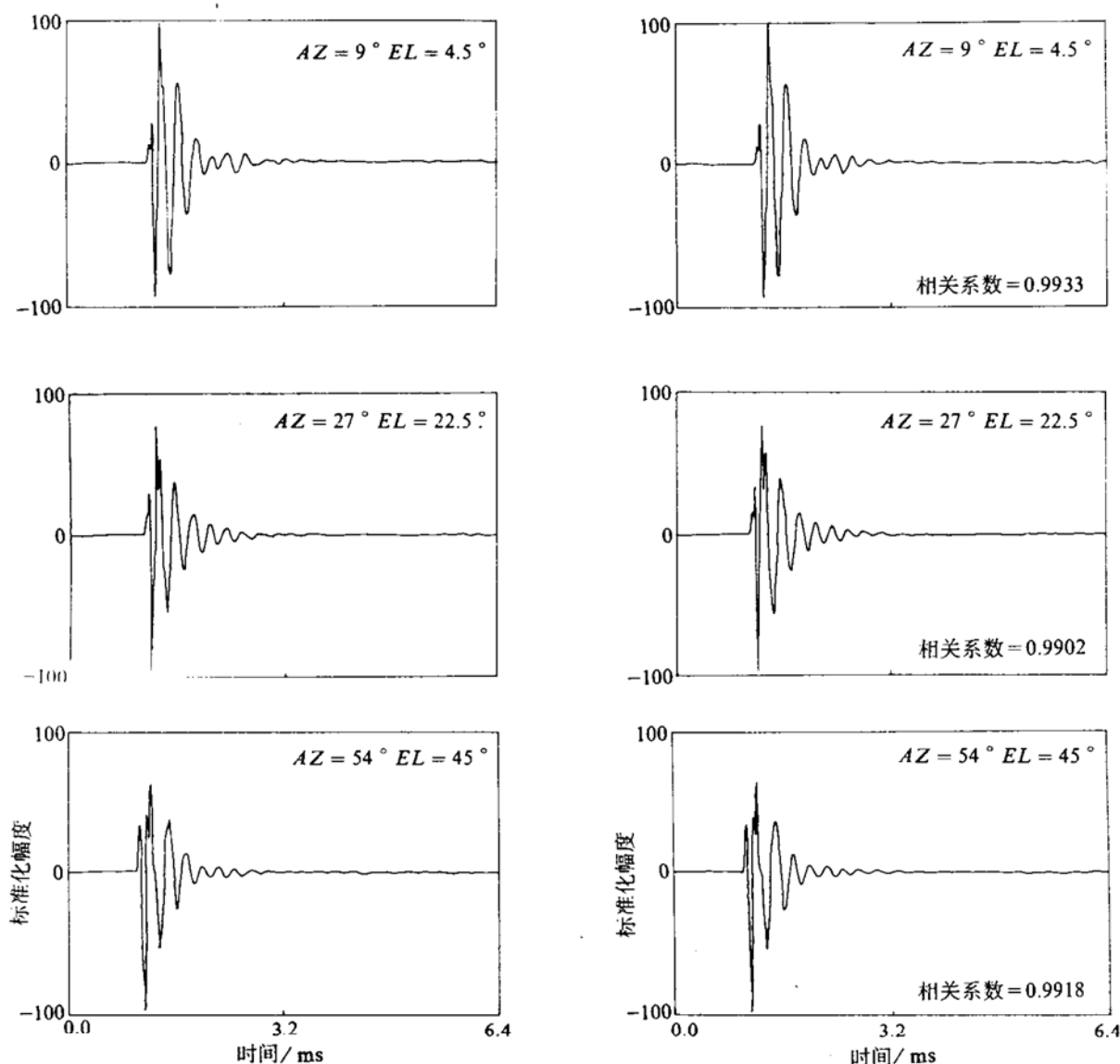


图 8 实际记录的听觉信号 (左边一列) 与 LMS 算法估计出的 FETF 所恢复出的信号 (右边一列) 相比较。

三、测量方法的改进

从上节讨论中我们可以看出, 经典的 FFT 方法估计 FETF 的缺点是明显的, 它将测量噪声当作听觉信号的一部分去估计 FETF。对于 LMS 方法它能克服这一缺点。但是在我们所关心的 2 kHz~30 kHz 范围内, 由于高频信号分量太弱, 由 LMS 方法估计出的 FETF 传递函数在高频部分也不够真实。图 9 给出了三个方向上的两种估计方法在 2 kHz~30 kHz 一段范围内的 FETF 幅频特性。不难看出, LMS 算法所估计的 FETF, 其幅度在 20 kHz~30 kHz 一段范围中明显地下降。图 8 所示的恢复信号没有很大失真是因为原始信号的高频分量本身就很弱。目前, 我们所用的单脉冲激励信号, 其幅度谱为 $\sin\omega/\omega$ 的形式, 从低频到高频是逐渐下降的, 其第一个零点频率位于脉冲宽度的倒数处。为保证激励信号具有一定的带宽, 其脉冲宽度总是非常窄

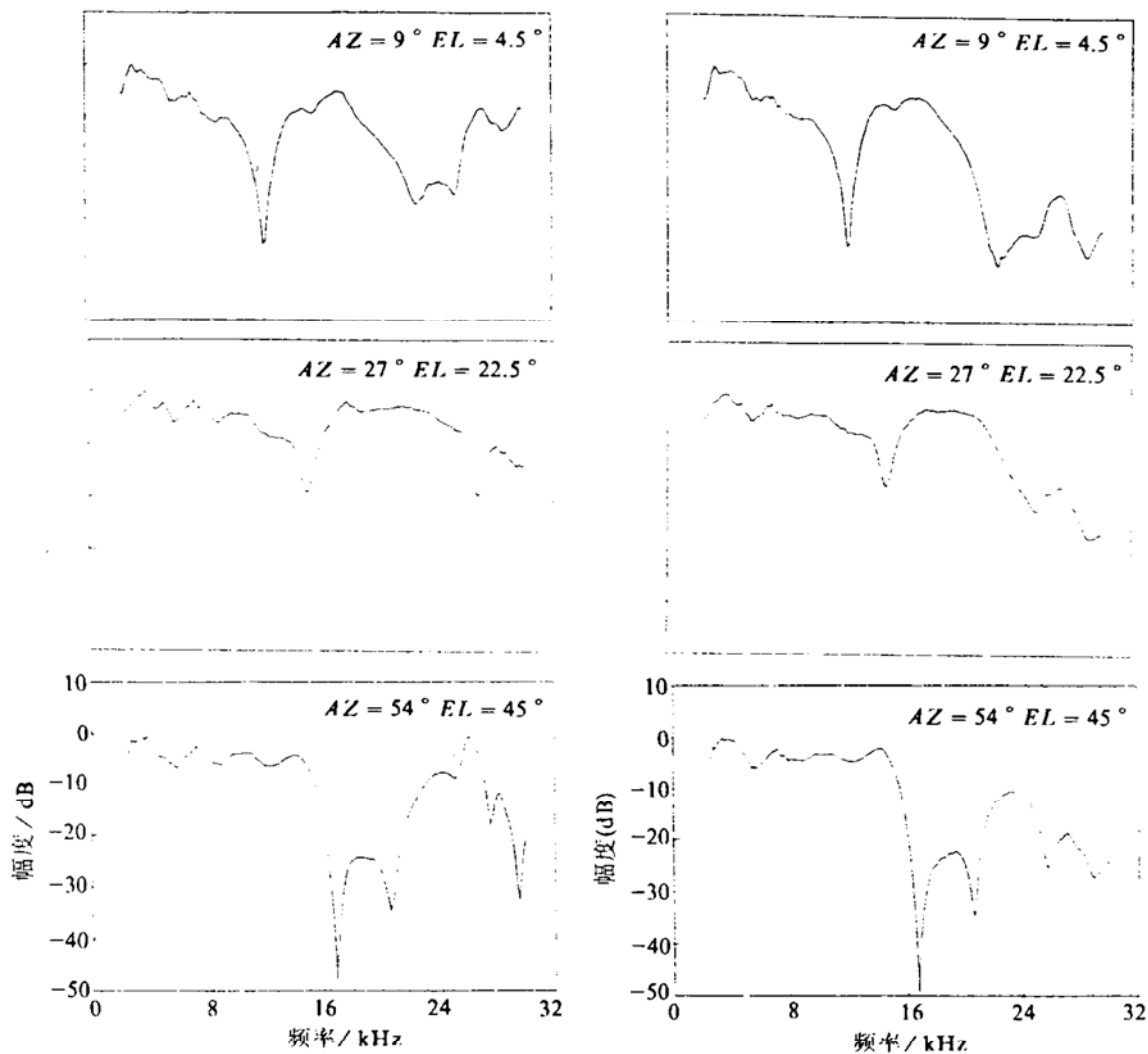


图9 在 2 kHz~30 kHz 范围内, 经典方法估计的 FETF (左边一列) 与 LMS 算法估计出的 FETF (右边一列) 相比较。

的, 通常为 $10\ \mu\text{s}$, 其能量便很有限。加上扬声器在 2 kHz~30 kHz 范围内的传递特性呈下降趋势, 自由声场激励信号在高频处的分量就更弱了。我们可以用两种方法提高自由声场激励信号的高频能量或整个信号的能量:

1. 改进测量信号使能量更集中于测试带宽内, 并对扬声器的非均匀传递特性进行补偿。测试信号可以用图 10(a) 所示的数字信号, 该信号的能量集中于 1 kHz~32 kHz 范围内, 通带特性也比较平坦, 见图 10(b)。为了保证自由声场激励信号接近图 10 所示信号的频率特性, 应该对扬声器的传递特性进行补偿。补偿办法是用图 10(a) 所示的信号与扬声器的反传递函数作卷积, 用此结果作为 D/A 变换器的输入信号 $d(n)$, 见图 1。扬声器的反传递函数可看作一有限冲激响应 (FIR) 滤波器。用此方法得到的扬声器输出将接近图 10(a) 的信号。这一补偿方法我们已成功地用于发送真实听觉空间信号的闭路系统^[11], 用以补偿传声器的失真。

2. 采用 Golay 码^[12]。B. Zhou 等于 1992 年提出了用 Golay 码代替单脉冲信号的方法。长度为 k 的 Golay 码是由两个码序 $a(k)$ 和 $b(k)$ 组成的, 每个码序含有 k 个元素。Golay 码可以由初始值 a_1, b_1 通过迭代产生

$$\begin{aligned} a_1 &= \{ +1 \quad +1 \}, \\ b_1 &= \{ +1 \quad -1 \}, \end{aligned} \quad (5)$$

而

$$\begin{aligned} a_2 &= \{ a_1 \quad b_1 \}, \\ b_2 &= \{ a_1 \quad -b_1 \}, \end{aligned} \quad (6)$$

任何长度 $L = 2^N$ 的 Golay 码可用类似于 (6) 式的方法产生。Golay 码的重要特性是两个码序的功率谱的和等于常数 $2L$ ，与单脉冲激励信号相比，它有非常平坦的频谱特性和充足的信号能量，其能量是单脉冲激励信号的 $2L$ 倍。

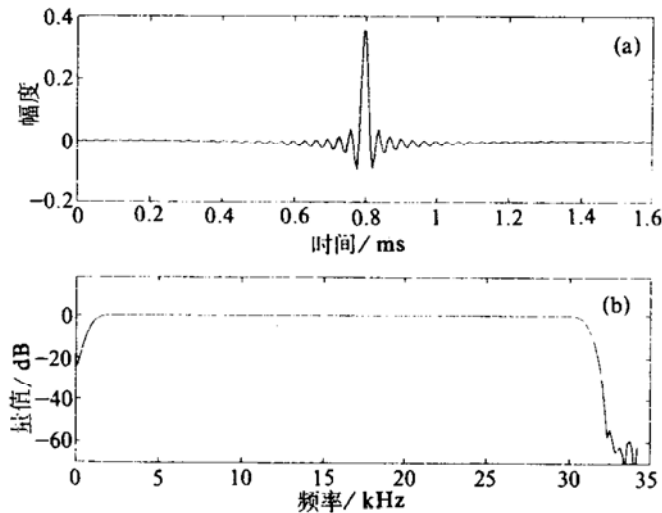


图 10 带通测试信号 (a) 时域波形 (b) 幅频特性

测量方法首先将 a_L 作为系统输入，得到系统输出，求得其付里叶变换为 $\text{FFT}(a_L)H(\omega)$ ；再将 b_L 作为系统输入，得到系统输出，求得其付里叶变换为 $\text{FFT}(b_L)H(\omega)$ 。最后用每一输出的付里叶变换乘上相应 Golay 码的付里叶变换的共轭值，并将此结果相加便得到了系统函数，即

$$H(\omega) = \{ \text{FFT}(a_L)H(\omega) \} \text{FFT}^*(a_L) + \{ \text{FFT}(b_L)H(\omega) \} \text{FFT}^*(b_L) / 2L \quad (7)$$

其中 $\text{FFT}^*(a_L)$ 和 $\text{FFT}^*(b_L)$ 分别为 $\text{FFT}(a_L)$ 和 $\text{FFT}(b_L)$ 的复共轭。

用 Golay 码测量时应注意的问题是测量所得的传递函数实质上包括了扬声器和探针式微音器的传递函数。这些在估计 FETF 时均应当从 $H(\omega)$ 中去除。另外，在实际估计 FETF 的过程中我们要对采集的自由声场信号和听觉信号做些预处理，以便消除其中的回波分量^[5, 13]。在采用 Golay 码测量时，其回波分量不象单脉冲激励信号那样易于区分，有必要研究一种新的回波抵消方法。

此外，伪随机序列或最大长度序列也具有类似于 Golay 码的特性^[14]。

四、结束语

本文提出的 FETF 估计方法已用于建立 VAS 模型，并在美国 Wisconsin 大学 Madison 分校的神经生理系的 VAX Alpha 计算机工作站以及 DSS (Digital Stimulus System) 系统上实现了这一模型^[13]。这一模型正用于听觉生理实验。

以上的全部工作是在美国 Wisconsin 大学 Madison 分校的神经生理系完成的。我们感谢该系的系主任 J. Brugge 博士对我们工作的支持，感谢 J. Hind 博士对我们工作的指导，并感谢

R. Kochhar 先生为我们提供了有关数据和计算机程序。这些工作也得到美国政府 NIH (National Institutes of Health) 基金的支持。

参 考 文 献

- [1] Stevens S S and Newman E B. The localization of actual sources of sound, *Am. J. Psychol.*, 1936, **48**: 297~306.
- [2] Bulter R A. The influence of the external and middle ear on auditory discrimination, in *Handbook of sensory physiology, Auditory System*, edited by Keidel W D and Neff W D, **2**, Springer-Verlag, Berlin, 1975.
- [1] Wightman F L and Kistler D J. Headphone simulation of free-field listening. I: Stimulus synthesis, *J. Acoust. Soc. Am.*, 1989, **85**: 858~867.
- [4] Poon P W F and Brugge J F. Virtual space receptive field of single auditory nerve fibers, *J. Neurophysiol.*, 1993, **70**: 667~676.
- [5] Chen J. Auditory space modeling and virtual auditory environment simulation, PhD thesis, University of Wisconsin-Madison, 1992.
- [6] Wenzel E M, Stone P K, Fisher S S and Foster S H. A system for three-dimensional acoustic visualization in a virtual environment workstation, *In Proceedings of the IEEE visualization*, 1990, San Francisco: 329~337.
- [7] Wenzel E M. Issues in the development of virtual acoustic environments, *J. Acoust. Soc. Am.*, 1992, **92**: 2332.
- [8] Musicant A D, Chan J C K and Hind J E. Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons, *J. Acoust. Soc. Am.*, 1990, **87**(2): 757~781.
- [9] Haykin S. Adaptive filter theory, Prentice-Hall, Englewood Cliffs, 1986.
- [10] Widrow B and Sterns S D. Adaptive signal processing, Prentice-Hall, Englewood Cliffs, 1985.
- [11] Jiashu Chen, Zhenyang Wu and Richard A Reale. Applications of least-squares FIR filters to virtual acoustic space, *Hear. Res.*, 1994, **80**: 153~166.
- [12] Zhou B, Green D M and Middlebrooks J C. Characterization of external ear impulse responses using Golay code, *J. Acoust. Soc. Am.*, 1992, **92**(2): 1169~1171.
- [13] Zhenyang Wu. FETF modeling implementation in the sound delivery-measuring system, Technical Report, University of Wisconsin-Madison, 1994.
- [14] MacWilliams F J and Sloane N J. Pseudo-random sequences and arrays, *Proceedings of the IEEE*, 1976, **64**(12): 1715~1729.