

早期衰变时间与信噪比对语言清晰度影响实验

魏卫刚¹, 葛剑敏¹, 张 华², 周兆驹³

(1. 同济大学 声学研究所, 上海 200092; 2. 北京市耳鼻喉科研究所, 北京 100005; 3. 山东建筑大学 建筑城规学院, 山东 济南 250001)

摘要: 随着计算机科学的发展, 声学的可听化技术被越来越广泛地应用到主观实验研究中. 声学软件 ODEON 可以建立很好的仿真声场, 利用 ODEON 的可听化技术, 对矩形高大空间和狭长空间的语言清晰度进行实验研究. 结果表明, 对于同一空间中的不同接收点的语言清晰度, 早期衰变时间的影响比信噪比的影响显著; 在语言清晰度为主要设计指标的厅堂中, 应该主要控制早期衰变时间, 使其不能超过 1.5 s.

关键词: 早期衰变时间; 信噪比; 语言清晰度; 声学软件
中图分类号: TU 112.2 **文献标识码:** A

A Comparison of Speech Intelligibility Between Different Parameters: Early-decay-time and Signal-noise-ratio

WEI Weigang¹, GE Jianmin¹, ZHANG Hua², ZHOU Zhaoju³
(1. Institute of Acoustics, Tongji University, Shanghai 200092, China; 2. Beijing Institute of Otolaryngology, Beijing 100005, China; 3. College of Architecture and Urban Planning, Shandong Jianzhu University, Ji'nan 250001, China)

Abstract: As the development of computer technology, auralization method is widely conducted in subjective experiment. ODEON, an acoustic software which can simulate the sound field nicely, was used as a tool to compare the speech intelligibility between large-volume space and long space. The results indicate that in the same space the early-decay-time performs more prominently than signal-noise-ratio. As a result, the early-decay-time is proposed to be designed as the main parameter and be less than 1.5 seconds in rooms where the main target is to obtain a high speech intelligibility.

Key words: early-decay-time; signal-noise-ratio; speech-intelligibility; acoustic software

来越广泛地应用到实际工程设计中. 这就非常便于在虚拟条件下研究建筑物内部的声场特性, 不同的声学参量对主观听感的影响. 不少商用的声学分析软件, 如 ODEON^[1-2]、EASE 等取得了较快发展, 虽然他们还存在一些不足^[3], 但是都可以胜任这些工作.

一些利用声场模拟进行声学实验的研究也相继展开, 有人撰文指出声场计算机仿真在设计中的可行性与优越性^[4], 借助声场仿真技术研究语言清晰度在不同的声场环境中的评价与差异^[5-7]. 也有人分析了声场模拟技术在实际工程中的应用情况^[8-9]. 不少学者研究了信噪比及声压级对教室语言清晰度的影响^[10-11], 也有人研究了不同的干扰噪声对语言清晰度的影响^[12]. 但是在同一房间中, 早期衰变时间更易受到其所在位置的影响, 在某些早期衰变时间较长的位置, 即使信噪比较大, 也不能获得良好的清晰度得分. 因此对比研究不同房间体型条件下的语言清晰度还是有必要的. 虽然这种借助于计算机仿真手段的研究不能完全跟实际情况相同, 却可以很方便地更改实验条件, 达到对比试验的目的.

本文将以高大空间和狭长空间作为研究对象, 利用声学软件 ODEON 8.5 建模仿真, 分析早期衰变时间与信噪比对语言清晰度的影响, 利用方差分析的方法, 确定显著影响语言清晰度的因素.

1 实验准备

1.1 选取研究对象

为了便于比较, 选取两个典型室内空间作为研究对象. 一个是以某体育训练馆为原型的高大空间为模型一, 见图 1; 另一个是某集体宿舍走廊为原型的狭长空间为模型二, 见图 2. 两个研究对象的空间形式混响时间等差别均比较大. 为了区分直达声为主的区域和混响声为主的区域, 将两个接收点分别

随着计算机科学的发展, 声学的可听化技术被越

收稿日期: 2009-05-04

作者简介: 魏卫刚(1980—), 男, 博士生, 主要研究方向为环境声学, 建筑声学. E-mail: sedwei@gmail.com

葛剑敏(1963—), 男, 教授, 工学博士, 博士生导师, 主要研究方向为环境声学. E-mail: gejianmin123@yahoo.com.cn

置小于混响半径和远大于混响半径的位置. 对于模型一,接收点1距离声源2 m,小于混响半径;接收点2距离声源16 m,大于混响半径. 对模型二,接收点1距离声源2 m,小于混响半径;接收点2距离声源8 m,大于混响半径.

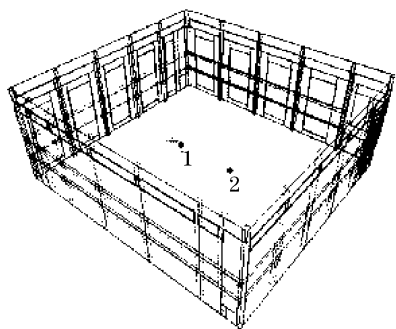


图1 模型一及接收点位置

Fig.1 Model 1 and receivers' position

1.2 建立模型获取双耳房间脉冲响应

首先在计算机辅助设计 CAD 软件中建立模型,

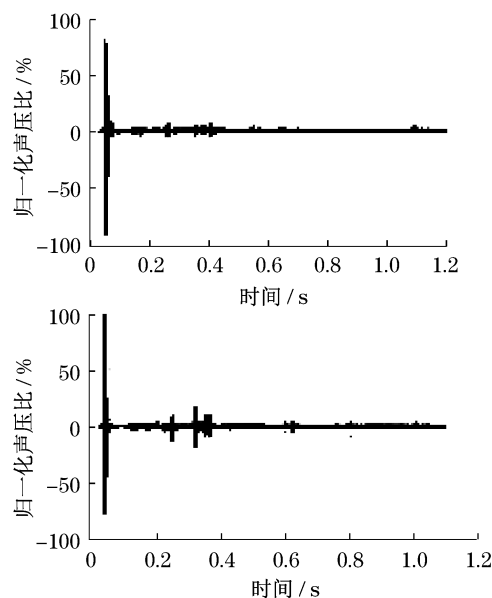


图3 模型一1#点 BRIR

Fig.3 BRIR of receiver 1, Model 1

1.3 录制测听信号

根据语言清晰度测试信号的要求^[14]录音采用汉语清晰度测试音节表——KXY 表的内容. 本次录音采用随机组合的方式,预先制作了20组测试音节表,每张表格都至少使用了奇偶两张KXY表中的音节. 采用男女发音人各一名,能讲正确的普通话,没

然后导入 ODEON. 由于软件的局限和声传播特性等原因,在建立模型时应根据实际情况进行必要的简化,不用建立过多的比较小的面. 这种适当的简化模型反而能够得到更好的计算结果^[13]. 由于选取的两个研究对象是实体对象,所以严格按照实际建筑内部壁面材料的吸声情况赋予模型材质,使得仿真声场尽量接近实际情况. 建好模型后,就可以得到双耳房间脉冲响应 (BRIR). 图3,图4是模型一两个接收点的双耳房间脉冲响应. 图5,图6是模型二的两个接收点的双耳房间脉冲响应. 利用 BRIR 与没有混响的信号卷积,然后再给定信噪比,就得到实验所需的信号了.

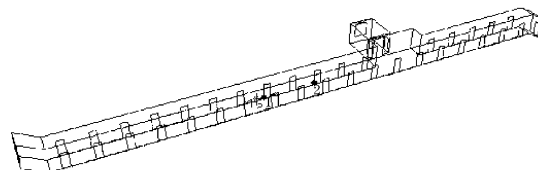


图2 模型二及接收点位置

Fig.2 Model 2 and receivers' position

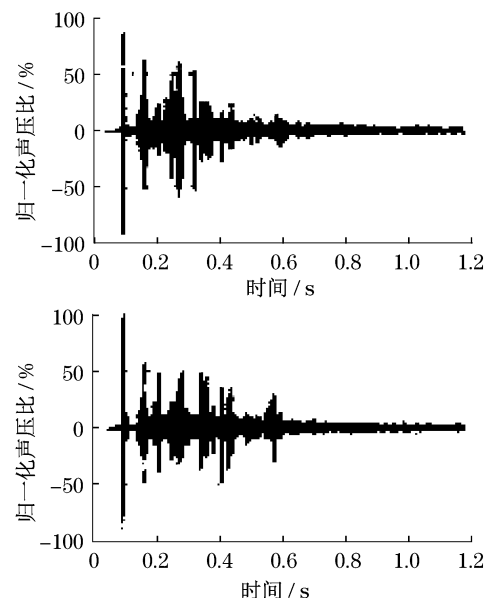


图4 模型一2#点 BRIR

Fig.4 BRIR of Receiver 2, Model 1

有语言缺陷. 整个录音过程在消声室内进行. 发音人以正常的讲话速度(约 $4 \text{ 音节} \cdot \text{s}^{-1}$),自然平稳地发音. 每组音节之间做适当停顿,记录完成后,将声音转化成 *.wav 格式文件,利用 Cool Edit Pro 软件将每组测听表之间的间隔准确调整在 5 s,便于对以后测试记录.

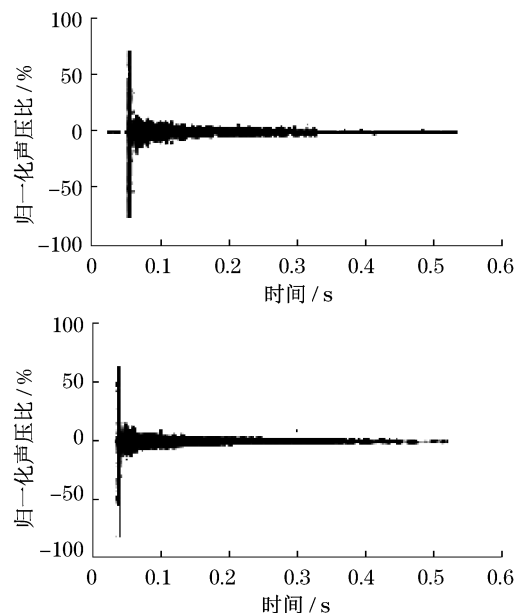


图5 模型二1#点 BRIR

Fig.5 BRIR of receiver 1, Model 2

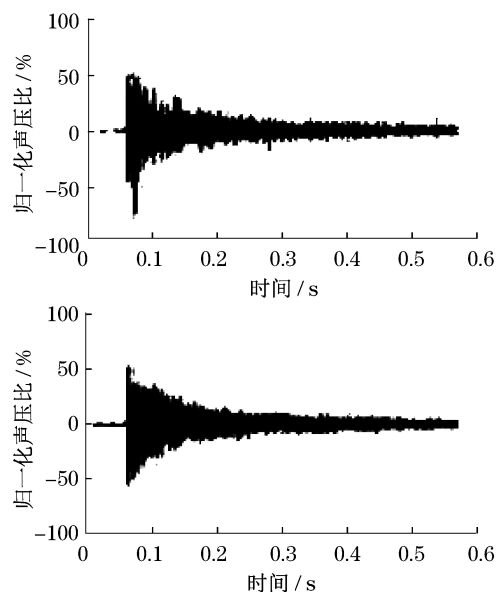


图6 模型二2#点 BRIR

Fig.6 BRIR of receiver 2, Model 2

1.4 测听信号与房间接收点处的 BRIR 卷积

利用 ODEON 8.5 自带的信号卷积功能实现听音信号的卷积. 为了防止测听人员在听到相同声音信号时, 由于本身的学习性而对实验结果产生干扰, 故利用所选的测试信号具有较高等价性的特点, 对于后续准备进行对比实验的测试组, 都采用不同的声音信号与双耳房间脉冲响应进行卷积. 为了便于识别卷积信号在厅堂中的位置, 对于每一个接收点都单独建立了一个任务.

2 可听化实验与结果分析

2.1 可听化实验

实验采用耳机重放的方式进行. 分别对模型一和模型二1#接收点按照信噪比为-10, -5, 0, 5 和 10 dB 进行打分, 对于2#接收点由于其在低信噪比时无法听清任何音节, 所以取消了几组低信噪比的测试, 增加了无干扰的清晰度实验, 并打分. 根据《声学语言清晰度的测试方法》提供的得分算法, 计算每张表的错误音节数 E_s , 则正确响应数 $N_s = 75 - E_s$, 其中, N_s 为答对问题数, E_s 为答错问题数, 75 为总问题数. 从而该记录表的音节清晰度得分为

$$S_i = \frac{N_s}{75} \times 100$$

如果有 n 位听音人参加, 则平均清晰度得分为

$$S = \sum_{i=1}^n S_i / n$$

2.2 结果分析

2.2.1 两个模型各点的清晰度得分

图7是模型一1#点和2#点的清晰度得分曲线. 图8是模型二1#点和2#点的清晰度得分曲线. 从图中可以看出, 对于模型一来说, 在信噪比相同的条件下, 1#点的清晰度明显优于2#点. 信噪比降低时1#、2#点清晰度都变差, 2#点下降明显. 对于模型二来说, 1#点的清晰度与2#点虽然听音者之间是有较大差异, 但是得分平均值无显著差别. 见各信噪比条件下模型各点对应的清晰度得分平均值比较表(见表1).

表1 不同信噪比模型各点对应的清晰度得分平均值

Tab.1 Mean intelligibility score of the two models

信噪比	-10	-5	0	5	10	无干扰
模型一1#点	3	27	43	56	69	
模型一2#点				8	28	57
模型二1#点	1	21	39	57	73	
模型二2#点		19	40	61	69	79

分析知道, 在确定信噪比时, 有用信号包含了房间信息, 即包含了接收点位置的混响信息, 所以对于2#点来讲, 音节本身的混响成了清晰度的干扰因素, 从两个接收点的脉冲响应图中也可以看出这一特征(见图3, 图4). 而1#点与声源的距离小于混响半径, 直达声是主要的, 混响对清晰度得分的影响就小, 所以对于相同信噪比, 1#点的清晰度要优于2#点. 对于模型二, 由于其混响时间较短, 音节本身的

混响对清晰度的干扰较小(见图5,图6),所以1#、2#点的清晰度在不同信噪比下的得分差异不大.对于相同编号的接收点,模型1和模型2的清晰度得分是有差异的(见图7,图8).但是不同接收点差异的显著性有很大差别.对于1#接收点来讲,两个模型的清晰度得分平均值差异很小(见图9),基本趋势与信噪比成线性增长关系.但是两个模型的混响时间相差较大(见图10).由此可以看出,两个模型混响时间的差异并没有影响1#点的清晰度得分.当信噪比为10 dB时,清晰度得分均超过了70%,可以想到如果音节在衰减10 dB前没有受到包括来自外界的噪声干扰、自身的混响干扰,那么此位置的语言清晰度得分就较高.可以看出声音的早期衰变是影响清晰度的重要因素^[15].

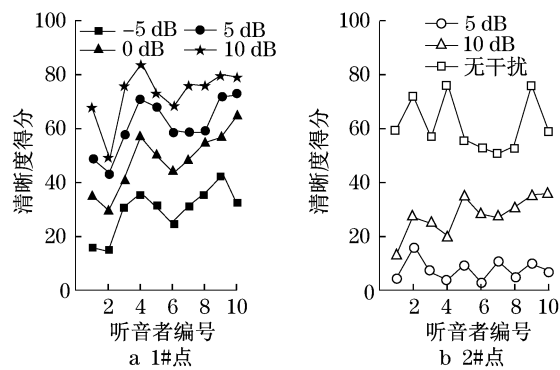


图7 模型一清晰度得分曲线

Fig.7 Intelligibility score in position 1 and position 2 of model 1

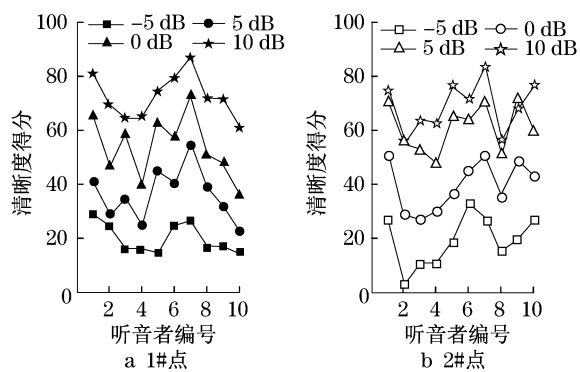


图8 模型二清晰度得分曲线

Fig.8 Intelligibility score in position 1 and position 2 of model 2

2.2.2 早期衰变时间和信噪比对清晰度影响的显著性对比

由于其他实验条件是固定的,故查看早期衰变时间的差异和信噪比的差异哪个因素对清晰度得分影响显著的问题,就变成了双因素方差分析的问题.

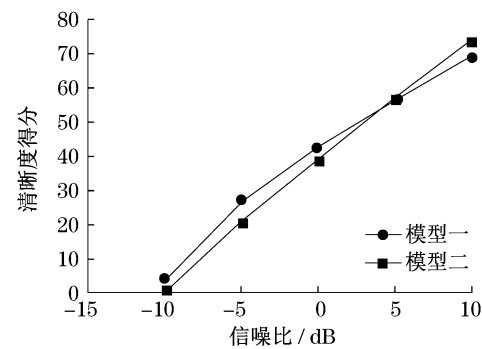


图9 两个模型1#点平均清晰度得分比较

Fig.9 Comparison of the mean articulation score between the two models

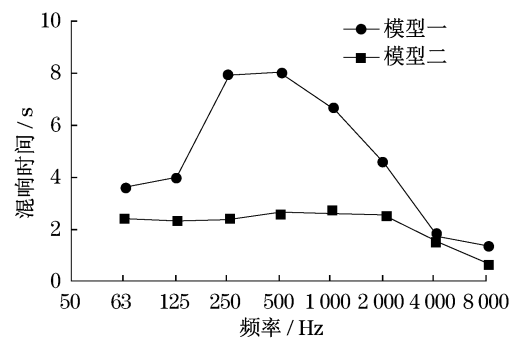


图10 两个模型1#点混响时间比较

Fig.10 Comparison of the reverberation time of position 1 between the two models

要对比的实验指标是1#、2#点清晰度的得分差异;影响因素A为1#、2#点早期衰变时间之差;影响因素B为5个不同的信噪比;影响因素A有 $r=2$ 个水平,一个是模型一的早期衰变时间(EDT)之差,另一个是模型二的EDT之差;影响因素B有 $s=5$ 个水平,即五个信噪比(SNR,其变量表示为 S_{NR} , $S_{NR} = -10, -5, 0, 5, 10$ dB).则可以简化成如下的统计模型,见表2.

表2 EDT与SNR对清晰度得分显著性影响模型

Tab.2 Remarkable analysis of the influence between EDT and SNR on articulation

		S_{NR}/dB				
		-10	-5	0	5	10
清晰度	模型1 EDT 差	0.033	0.273	0.434	0.482	0.435
得分差值/%	模型2 EDT 差	0.005	0.029	0	0.027	0.060

注:表内数据为模型1、模型2对应信噪比下,1#、2#点清晰度得分百分比的差值.

由分析知,信噪比与EDT互不相干,故满足无交互作用的条件.上述模型就成了无交互作用的双因素方差分析.则需要的参数计算如下^[16]:

$$S_T = \sum_i \sum_j x_{ij}^2 - \frac{1}{n} (\sum_i \sum_j x_{ij})^2, \quad n = rs$$
$$S_A = \frac{1}{s} \sum_i (\sum_j x_{ij})^2 - \frac{1}{n} (\sum_i \sum_j x_{ij})^2$$
$$S_B = \frac{1}{r} \sum_j (\sum_i x_{ij})^2 - \frac{1}{n} (\sum_i \sum_j x_{ij})^2$$
$$S_e = S_T - S_A - S_B$$

式中: S_T 为总偏差平方和; S_A, S_B 为组间平方和; S_e 为组内误差; x_{ij} 为不同 EDT 或 SNR 对应的样本值; r 为 EDT 差的因素水平取值 2; s 为信噪比的因素水

平取值 5. $\frac{1}{n} (\sum_i \sum_j x_{ij})^2 = 0.432$, 经计算得, $S_T = 0.369$; $S_A = 0.337$; $S_B = 0.017$; $S_e = 0.014 5$; 其中, F 为 F 分布在不同自由度下的值. 当 $F > F_{0.01}$ 时, 影响因素与试验指标高度相关, 当 $F \leq F_{0.1}$ 时, 影响因素对试验指标无显著影响, 计算结果见表 3. 可见, 在混响场中 EDT 与 SNR 相比, 对清晰度的影响更显著. 对于模型 1 的 2# 点来讲, 其 EDT 太长, 即使给定很高的信噪比也不能提升相应的清晰度得分.

表 3 双因素方差分析计算表

Tab.2 Two-factor analysis of variance about articulation						
方差来源	平方和	自由度	均方	F 值	临界值	显著性
EDT 之差	$S_A = 0.337$	$r - 1 = 1$	$\bar{S}_A = 0.337$	$F = \bar{S}_A / \bar{S}_e = 102.01$	$F_{0.01}(1, 4) = 21.2$	高度显著
不同 SNR	$S_B = 0.017$	$s - 1 = 4$	$\bar{S}_B = 0.004 28$	$F = \bar{S}_B / \bar{S}_e = 1.185 5$	$F_{0.01}(4, 4) = 16,$ $F_{0.1}(4, 4) = 4.11$	无显著影响
误差	$S_e = 0.014 5$	$(r - 1) \cdot (s - 1) = 4$	$\bar{S}_e = 0.003 61$			
总和	0.369	9				

由于实验模拟实际情况, 所以 1# 点和 2# 点不受背景噪声的干扰. 在近声源点直达声声压级高, 反射声声压级相对较低, 反射声就没有参与前期的能量衰减过程, 则接收点能量衰减 10 dB 所用时间就短, 相应的 EDT 就短; 在远声源点, 则情况不同, 直达声本身在传递过程中就有衰减, 到达接收点时已经有衰减, 如果壁面吸声较弱的话, 则反射声到达接收点时就相对较强, 它的能量值如果大于接收点直达声衰减 10 dB 以后的值, 则其 EDT 就会大大加长, 清晰度就会明显下降. 从两个模型脉冲响应上也能看到这个特点(见图 3~图 6).

3 确定保证清晰度的 EDT 临界值

根据上面的实验分析, 如果声源衰减 10 dB 没有受到下一个讲话音节、周围壁面的反射、背景噪声等因素的影响, 则清晰度得分较高, 如两个模型 1# 点的清晰度得分, 而 EDT 又定义为声源停止发声后声场衰减前 10 dB 的斜率所确定的时间的 6 倍. 由此出发可以计算处保证清晰度的 EDT 临界值. 统计发现正常讲话的语速为 $4 \text{ 音节} \cdot \text{s}^{-1}$, 所以两个音节发声间隔为 250 ms, 也就是声源在此时间内衰减完成 10 dB, 则能保证很高的清晰度. 显然, 由此出发的处 EDT 的临界值为 $\text{EDT}_{\max} = 250 \times 6 = 1\,500 \text{ ms}$. 即 EDT 的临界值是 1.5 s, 如果 EDT 大于 1.5 s, 则清晰度就会下降.

4 结论

结果表明, 同一空间中早期衰变时间对清晰度的影响比信噪比要显著; 在信噪比达到 10 dB 以后, 混响半径内清晰度得分在 70% 以上, 由此出发, 推算出保证清晰度的 EDT 临界值为 1.5 s.

参考文献:

[1] Rindel J H, Christensen C L. Room acoustic simulation and auralization-how close can we get to the real room? [C]//The Eighth Western Pacific Acoustics Conference. Melbourne: [s. n.], 2003:1-8.

[2] Ingolf Bork. A comparison of room simulation software-the 2nd round robin on room acoustical computer simulation[J]. Acta Acoustica, 2000(86):943.

[3] 曾向阳, Christensen C L, Rindel J H. 建筑声学软件 ODEON 低频性能分析[J]. 电声技术, 2005(3):10.
ZENG Xiangyang, Christensen C L, Rindel J H. Low frequency performance of the building acoustics software ODEON[J]. Audio Engineering, 2005(3):10.

[4] Naylar G M. ODEON——another hybrid room acoustical model [J]. Applied Acoustics, 1993(38):131.

[5] 彭健新, 吴硕贤. 厅堂声场计算机仿真[J]. 南方建筑, 2003(2):53.
PENG Jianxin, WU Shuoxian. Sound field simulation of buildings by the tool of computer[J]. South Architecture, 2003(2):53.

[6] 彭健新. 声场仿真在室内语言清晰度评价中的应用[J]. 电声技术, 2002(11):16.
PENG Jianxin. Application of sound field emulation in the evaluation of indoor speech intelligibility [J]. Audio Engineering, 2002(11):16.