

# 一种基于物理特性和感知特性的混响模拟方法

徐华兴, 夏日升, 李军锋\*, 颜永红

中国科学院声学研究所, 北京 100190

\* 通信作者. E-mail: lijunfeng@hccl.ioa.ac.cn

收稿日期: 2014-10-31; 接受日期: 2015-04-06; 网络出版日期: 2015-05-20

中国科学院战略性先导科技专项 (批准号: XDA06030100) 和国家自然科学基金 (批准号: 11461141004) 资助项目

**摘要** 人工混响在许多领域 (如现场音乐和计算机游戏) 具有重要应用前景. 传统的基于物理特性和感知特性的人工混响模拟方法分别具有计算量大和缺少房间具体特征等缺点. 本论文提出了一种混合人工混响 (hybrid physically- and perceptually-based reverberation, HPPR) 模拟方法, 首先利用基于物理特性的镜像法产生房间脉冲响应, 通过与“干”信号进行卷积模拟产生早期混响; 然后利用基于感知特性的反馈延时网络模拟产生后期混响. 为了将模拟产生的早期混响和后期混响更好地结合起来, 提出了一种早期混响能量衰减的参数化预测方法, 据此实现反馈延时网络参数的估计, 保证了早期混响到后期混响在时-频域的平滑过渡. 数值实验和主观测听实验验证了所提混响模拟方法的有效性.

**关键词** 混响模拟 镜像法 反馈延时网络 能量衰减曲面 房间声学

## 1 引言

混响在我们的日常聆听体验中扮演着重要角色, 其人工模拟有着极其广泛的应用 (如音乐和计算机游戏). 人工混响的模拟方法分为 3 类: 基于物理特性的模拟方法、基于感知特性的模拟方法和混合混响模拟方法<sup>[1]</sup>. 基于物理特性的混响合成方法从物理角度出发精确模拟了房间混响特性, 但其巨大的计算量限制了在实时系统中的运用, 如交互式声环境重放<sup>[2,3]</sup>. 基于感知特性的混响模拟方法, 如反馈延时网络 (feedback delay network, FDN), 仅模拟混响中对其听觉感知重要的部分, 降低了运算量, 提高了计算速度<sup>[4~6]</sup>. 虽然 FDN 能够模拟产生出高质量的混响感知效果, 但不能很好地反映待模拟环境的声学特性<sup>[7,8]</sup>.

考虑到传统的基于物理特性和感知特性的混响模拟方法的优缺点, 近年来, 通过综合利用这两种方法的混合人工混响模拟方法得到了广泛关注<sup>[9,10]</sup>. 人工混响混合模拟方法的基本依据是混响可分解成早期反射 (包括直达声) 和后期混响. 学者们提出了一些计算高效同时又不影响混响听觉感知效果的混合混响模拟方法<sup>[9,10]</sup>. Rindel<sup>[9]</sup> 利用镜像法 (image-source model, ISM) 建模早期反射, 后期混响采用声线跟踪方法 (ray-tracing method, RTM) 实现. 此外, Murphy 和 Stewart<sup>[10]</sup> 提出从实际测量的房间脉冲响应 (room impulse response, RIR) 中直接截取一个短的响应 (finite impulse response, FIR), 然后通过卷积运算产生早期混响, 通过快速的 FDN 模拟后期混响. 然而所有这些混合混响模拟

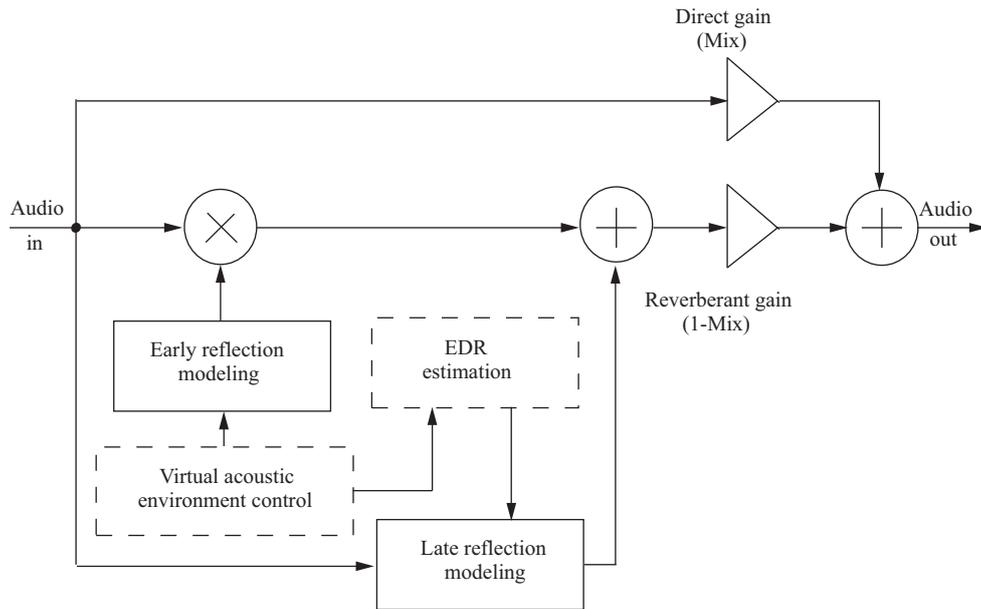


图 1 HPPR 混响模拟方法结构框图

Figure 1 Block diagram of the proposed HPPR reverberation approach

方法都没有明确的机制确保由不同方法模拟的早期混响和后期混响之间的平滑过渡. 为了解决这个问题, Lehamann 等<sup>[11]</sup>利用根据 ISM 方法建模的早期混响能量衰减曲线 (energy decay curve, EDC) 确定后期混响 (衰减随机噪声) 的能量衰减包络. Primavera 等<sup>[12]</sup>根据模拟的早期反射, 采用基于同步扰动随机近似法 (simultaneous perturbation stochastic approximation, SPSA) 计算用于后期混响模拟方法中的滤波器参数. 值得注意的是, 在这些方法中, 从早期混响到后期混响在时 — 频域的平滑过渡仍未得到明确解决; 此外, 它们仍需要从早期反射中离线计算得到预先模拟的 RIR 和后期混响模拟参数<sup>[12,13]</sup>.

本论文提出一种基于物理特性和感知特性的混响 (hybrid physically- and perceptually-based reverberation, HPPR) 模拟方法, 其中早期反射由 ISM 建模产生的早期 RIR 卷积得到, 后期混响利用 FDN 实现. 为了使模拟的早期反射和后期混响很好地结合起来, 进一步提出一个参数化预测模型估计 ISM 建模产生的早期反射的能量衰减曲面 (energy decay relief, EDR), 据此 FDN 参数实时自动计算得到. 此处理过程确保了早期反射到后期混响在时 — 频域的平滑过渡, 同时模拟的后期混响能够一定程度上反映待模拟环境的声学特性. 最后通过数值实验和主观测听实验, 验证了 HPPR 混响模拟方法的有效性.

## 2 基于物理特性和感知特性的混响模拟方法

本节介绍所提出的 HPPR 混响模拟方法, 其主要包括: (1) 早期混响是由 ISM 建模得到的冲击响应卷积“干”信号得到; (2) 后期混响是由 FDN 实现; (3) 参数化预测 ISM 建模的早期混响的 EDR. 最后, 模拟产生的早期混响、后期混响以及直达声相加产生最终的完整混响. HPPR 混响模拟方法如图 1 所示, 其中图中虚线框表示所期望模拟的房间参数, 然后据此设定早期反射及计算 EDR, 再利用 EDR 参数调整 FDN 参数; 实线框图表示输入的音频实际混响模拟处理流程.

## 2.1 早期混响模拟

HPPR 混响模拟方法中, 早期混响模拟是通过卷积 ISM 建模得到的 RIR 与“干”信号来实现. ISM 最初是由 Allen 和 Berkley<sup>[14]</sup> 提出的, 其基本思想是在一个封闭室内环境中, 一个点声源产生的混响可以由一系列特定位置的虚拟声源产生的声压和来模拟. 假设声源和接收者分别在位置  $P_s = [x_s, y_s, z_s]$  和  $P_r = [x_r, y_r, z_r]$ , 房间尺寸为  $r = [L_x, L_y, L_z]$ . 房间所有表面的声学特性用反射系数  $\beta_{x\xi}$ ,  $\beta_{y\xi}$  和  $\beta_{z\xi}$  来表征, 其中  $\xi=1, 2$ ,  $\xi=1$  表示紧邻坐标系原点的墙壁,  $\xi=2$  表示相反的墙壁. 传统的镜像法对房间早期反射的模拟主要利用房间的几何声学特性, 很少考虑空气对声音信号的吸收特性. 由于空气吸收的影响, 声音在传输过程中高频能量衰减通常比低频更快. 因此考虑了空气吸收影响之后, ISM 修正建模得到的 RIR 可以表示为

$$h(\omega) = \sum_{p \in \mathcal{P}} \sum_{m \in \mathcal{M}} \left\{ \beta_{x_1}^{|m_x - q|} \beta_{x_2}^{|m_x|} \beta_{y_1}^{|m_y - u|} \beta_{y_2}^{|m_y|} \beta_{z_1}^{|m_z - v|} \beta_{z_2}^{|m_z|} \right. \\ \left. \times \frac{\exp [j(\omega/c) |R_p + R_m|]}{4\pi |R_p + R_m|} \underbrace{\exp (-\alpha |R_p + R_m|)}_{\text{Air absorption}} \right\}, \quad (1)$$

其中,  $\omega = 2\pi f$  表示以弧度单位的频率,  $f$  表示以 Hz 为单位的频率,  $c$  表示声音的速度,  $\mathcal{P} = \{(q, u, v) : q, u, v \in [0, 1]\}$  表示一系列三维点  $(q, u, v)$  的集合.  $\mathcal{M} = \{(m_x, m_y, m_z) : -M \leq m_x, m_y, m_z \leq M\}$  表示一系列三维点  $(m_x, m_y, m_z)$  的集合, 其中  $M$  表示反射阶数.  $\alpha$  表示由于空气吸收引起的能量衰减因子, 它是频率和空气特性的函数 (如空气压力和温度等)<sup>[15]</sup>.  $R_p = [(1-2q)x_s - x, (1-2u)y_s - y, (1-2v)z_s - z]$  表示相对于坐标原点的声像位置,  $R_m = [2m_x L_x, 2m_y L_y, 2m_z L_z]$  是考虑高阶反射的虚拟房间尺寸.

## 2.2 后期混响模拟

利用 FDN 模拟后期混响最早由 Stautner 和 Puckette<sup>[7]</sup> 提出的, 随后 Jot 和 Chaigne<sup>[8]</sup> 对其进行了改进. FDN 由一系列延时通路经过反馈矩阵连接反馈环路组成. 各延时通路的长度 (秒) 由  $\tau_i = m_i T$  给出, 其中  $T = 1/F_s$  为抽样周期,  $m_i$  可由方法 [16] 计算所得, 反馈矩阵可以利用 Householder 矩阵实现<sup>[7]</sup>. 为进一步提高后期混响的自然度, 在各个延时通路中插入相应的低通滤波器来模拟混响随频率的衰减特性<sup>[8,17]</sup>. 低通滤波器的传输函数  $H_i$  可由下式表示:

$$20 \log H_i(\omega) = -60 \frac{\tau_i}{T_r(\omega)}, \quad (2)$$

很明显,  $H_i(\omega)$  由衰减时间  $\tau_i$  和混响时间  $T_r(\omega)$  计算所得, 其中  $T_r(\omega)$  表示相对于直达路径的能量衰减  $r$  dB 的混响时间. 由于  $T_r(\omega)$  是频率相关的, 并且可以从模拟出的早期混响中估计, 所以  $H_i(\omega)$  对从早期混响到后期混响在时—频域的平滑过渡产生重要影响. 因此, 有效地计算混响时间随频率增加而下降的曲线  $T_r(\omega)$  变得尤为重要.

## 2.3 参数化早期混响能量衰减预测方法

为了计算与频率相关的混响时间  $T_r(\omega)$ , 本小节提出一种早期混响 EDR 的参数化预测方法. EDR 的传统计算方法是通过在给定一个时间内, 测量或模拟的 RIR 每一频带的能量和<sup>[17]</sup>. EDR 传统计算方法需要耗费巨大的计算量, 为了减小计算量, 依据文献 [18] 中方法, ISM 模拟的早期混响的 EDR 可由待建模房间的参数决定. 文献 [18] 利用传统 ISM 方法建模早期反射 RIR 仅考虑房间墙壁对声音的

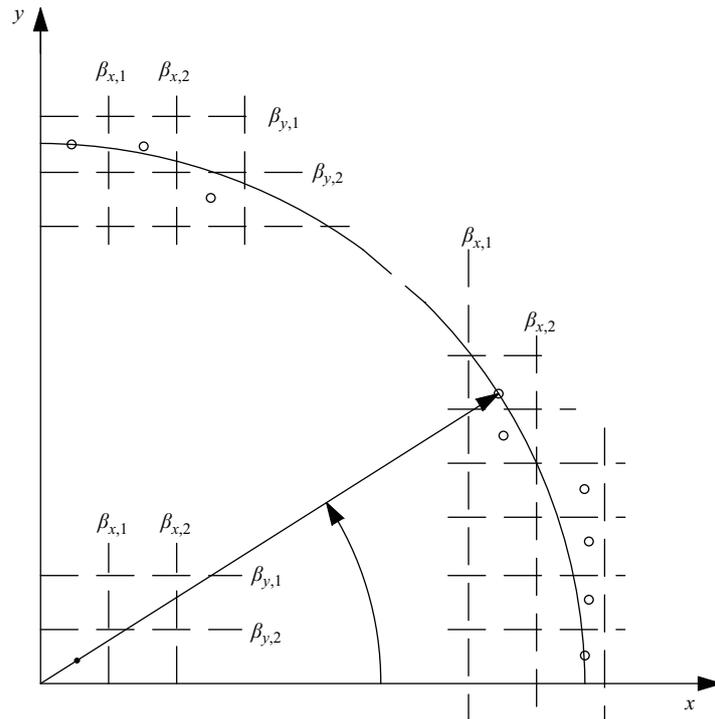


图 2 ISM 的二维示意图

**Figure 2** Two-dimensional representation of the ISM principle. The rectangular room enclosing the black point is the original room, circles (o) represent the considered image sources, the  $\beta$  parameters indicate the reflection coefficients of the corresponding boundaries

吸收作用, 提出了计算某时刻 RIR 所有频率上的 EDC 的预测方法. 本文在此基础上, 利用 ISM 建模早期反射 RIR 时引入与空气吸收特性相关的衰减因子, 考虑墙壁和空气在不同频率或频带对声音能量衰减的不同影响, 进而得到了在不同频带内由于墙壁和空气的吸收作用引起的声音随时间衰减的曲线即 EDR 的预测模型.

ISM 方法中每一个镜像源可视为沿接收方向以声速传播的声粒子 (sound particle), 然后在边界反射衰减吸收. 利用这种原理, 给定  $t$  时刻和频带  $\omega$ , 传声器接收到的声能  $h_p(t, \omega)$  对应所有位于 (或接近于) 传声器球面直径  $\rho = c \cdot t$  的所有镜像声源在接受点的声能之和. 先设定为二维平面, 具体如图 2 所示. 设  $\rho \gg (L_x + L_y)$ , 为简化分析假设传声器位于分析的坐标系中心, 以上假设带来的误差, 在球面直径  $\rho$  增大时可忽略. 考虑角度  $\theta$  方向的虚拟镜像声源, 在到达接收点之间, 经过  $x$  方向,  $y$  方向的镜像墙声源数目分别为:  $W_x \triangleq W_x(\rho, \theta)$  和  $W_y \triangleq W_y(\rho, \theta)$ . 因此, 接收点的声能可表示为

$$P(\rho, \theta, \omega) = \exp(-\alpha\rho) \frac{(\beta_{x,1}^2)^{W_x/2} (\beta_{x,2}^2)^{W_x/2} (\beta_{y,1}^2)^{W_y/2} (\beta_{y,2}^2)^{W_y/2}}{(4\pi\rho)^2}. \quad (3)$$

RIR 的功率参数估计  $h_p(t, \omega)$  可表示为

$$h_p(t, \omega) = \sum_{i \in \mathcal{N}} P(\rho, \theta_i, \omega), \quad (4)$$

其中,  $\mathcal{N}$  为考虑角度内镜像声源数目. 将上述等式利用 Riemann 和逼近, 其可表示为以角度  $\theta$  为变量

的连续函数积分, 二维情况下时 — 频域 RIR 的功率参数估计  $h_p(t, \omega)$  表示为

$$\hat{h}_p(t, \omega) = \frac{4\rho}{\bar{r}} \int_0^{\pi/2} P(\rho, \theta, \omega) d\theta, \tag{5}$$

其中,  $\bar{r} = (L_x + L_y) / 2$ .

将上述推导过程扩展到三维空间, 三维空间坐标位置的虚拟声源在接受点处的能量可以由 Green 函数求解 Helmholtz 方程所得, 离散求和还可视为 Riemann 和, 进而可以转化为一连续函数在球面上的积分. 三维空间下时 — 频域 RIR 的功率参数估计  $h_p(t, \omega)$  表示为

$$\hat{h}_p(t, \omega) = \frac{8\rho \exp(-\alpha\rho)}{\bar{r}} \int_0^{\pi/2} \int_0^{\pi/2} P(\rho, \theta, \phi, \omega) d\theta d\phi, \tag{6}$$

其中,  $\bar{r} = (L_x + L_y + L_z) / 3$ . 求解上述二重积分得到封闭形式的能量表达式如式 (7) 所示:

$$\hat{h}_p(t, \omega) = \frac{\exp(-\alpha\rho)}{8\rho\bar{r}} \begin{cases} \frac{B_z}{\log\left(\frac{B_y}{B_x}\right)} \left\{ \text{Ei}\left[\log\left(\frac{B_z}{B_x}\right)\right] + \log\left[\log\left(\frac{B_z}{B_x}\right)\right] - \text{Ei}\left[\log\left(\frac{B_z}{B_y}\right)\right] - \log\left[\log\left(\frac{B_z}{B_y}\right)\right] \right\}, & \text{如果 } B_x \neq B_y \neq B_z, \\ \frac{B_z}{\log\left(\frac{B_z}{B}\right)} \left\{ \text{Ei}\left[\log\left(\frac{B_z}{B}\right)\right] + \log\left[\log\left(\frac{B_z}{B}\right)\right] + \gamma \right\}, & \text{如果 } B_z = B_y \neq B_x \triangleq B \text{ or } B_z = B_x \neq B_y \triangleq B, \\ \frac{B - B_z}{\log\left(\frac{B}{B_z}\right)}, & \text{如果 } B_z \neq B_x = B_y \triangleq B, \\ B, & \text{如果 } B_x = B_y = B_z \triangleq B, \end{cases} \tag{7}$$

其中,

$$B_x = (\beta_{x,1}\beta_{x,2})^{\rho/L_x}, \tag{8}$$

$$B_y = (\beta_{y,1}\beta_{y,2})^{\rho/L_y}, \tag{9}$$

$$B_z = (\beta_{z,1}\beta_{z,2})^{\rho/L_z}, \tag{10}$$

$\gamma = 0.5772157$  是 Euler-Mascheroni 常数,  $\text{Ei}(\cdot)$  表示一阶指数积分. 值得说明的是, 在利用 ISM 建模 RIR 早期反射时引入了表征空气吸收特性的频率衰减因子, 因此式 (7) 与文献 [18] 中式 (32) 相比: RIR 功率估计表示为时 — 频域的函数, 一定程度上反映待模拟环境的声学特性. 利用 RIR 的功率参数估计  $\hat{h}_p(t, \omega)$ , ISM 模拟的早期反射的 EDR 可表述为

$$\hat{E}(t, \omega) = 10 \log_{10} \left( \int_t^\infty \hat{h}_p(\xi, \omega) d\xi \right). \tag{11}$$

### 2.4 混响时间估计

基于式 (11) 得到的 EDR 估计  $\hat{E}(t, \omega)$ , 频率相关的混响时间  $T_r(\omega)$  可以根据式 (2) 来确定低通滤波器频率响应  $H_i(\omega)$ . 本研究中, 混响时间估计用  $T_{20}(\omega)$  表征, 其定义为 RIR 在频带  $\omega$  处的能量从 -5 dB 衰减到 -25 dB 所需时间:

$$T_{20}(\omega) = \hat{E}^{-1}(-25, \omega) - \hat{E}^{-1}(-5, \omega), \tag{12}$$

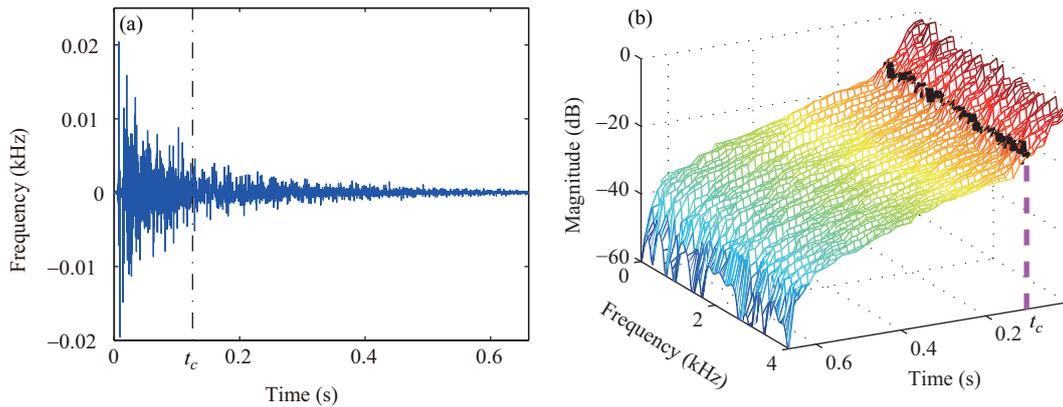


图 3 (网络版彩图) HPPR 模拟的房间 (6.0 m×5.0 m×3.0 m) 脉冲响应和能量衰减曲线

Figure 3 (Color online) Simulated by the proposed HPPR reverberation simulation approach in an environment with the dimension 6.0 m×5.0 m×3.0 m. (a) Room impulse response (RIR); (b) energy decay relief (EDR)

式中,  $\hat{E}^{-1}(\xi, \omega)$  对应时间滞后  $t_\xi$ , 其中  $E(t_\xi, \omega) = \xi$ .

HPPR 混响模拟方法, 其早期反射的模拟利用 ISM 建模只需建模很少阶数的 RIR, 减少了计算量; 后期混响利用 FDN 实现, 避免了耗时的卷积运算. 进一步的提出参数化预测模型估计 ISM 建模产生的早期反射 EDR, 基于此再实时设计 FDN 参数, 确保了早期反射和后期混响在时 — 频域的平滑过渡; 此外引入衰减项模拟声音传播的衰减特性来反映环境的声学特性.

### 3 实验与结果

为了验证 HPPR 混响模拟方法的有效性, 本节通过数值实验和主观测听实验来对比评价本文混合混响模拟方法 (HPPR)、标准 ISM 方法 (ISM) 以及改进的 ISM (ISM-DRM) 方法.

#### 3.1 数值模拟 RIR

为了验证 HPPR 混响模拟方法的有效性, 首先设定房间尺寸为  $6.0 \times 5.0 \times 3.0 \text{ m}^3$ , 根据 Sabin 公式<sup>[19]</sup> 计算得到其混响时间为 436 ms. 根据文献 [11] 中的建议, 利用 ISM 只对 RIR 能量衰减  $\Delta_c = 15 \text{ dB}$  之前的早期反射进行模拟, 对应的混合时间 (即早期混响和后期混响的分界时间点) 为  $t_c = 130 \text{ ms}$ .

HPPR 混响模拟方法得到的脉冲响应如图 3(a) 所示, 其中垂直虚线表示 ISM 模拟的早期混响和 FDN 模拟的后期混响分界点  $t_c$ . 从图 3(a) 可以看出 HPPR 模拟得出的早期混响和后期混响具有非常相似的包络, 这表明 ISM 模拟的早期混响和 FDN 模拟的后期混响在时域实现了平滑过渡. 图 3(b) 画出了利用 HPPR 模拟出的 RIR 在时 — 频域的 EDR, 从图中可以看到早期混响和后期混响在频域上的连续性. 由于在时频域早期混响和后期混响之间不存在任何不连续或者突变, 这表明 ISM 模拟得到的早期混响和 FDN 模拟得到的后期混响在时 — 频域实现了平滑过渡.

#### 3.2 主观测听实验

在主观测听实验中, 我们共模拟产生了 9 个不同房间的冲击响应, 房间体积  $V \in [30, 100] \text{ m}^3$  和混响时间  $T_{60} \in [0.2, 0.6] \text{ s}$  随机分布. 在每一个房间中, 声源位置和接收位置也是随机选择. 对于 ISM 和

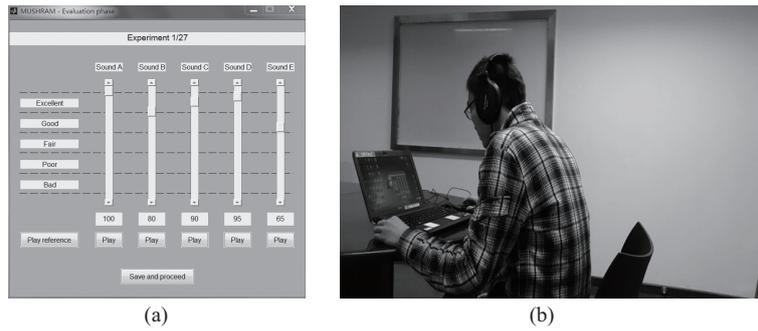


图 4 实验所用 Matlab 评测界面和听者受试环境

Figure 4 Matlab GUI and listener test environment for evaluating the algorithms. (a) Matlab GUI; (b) listener test environment

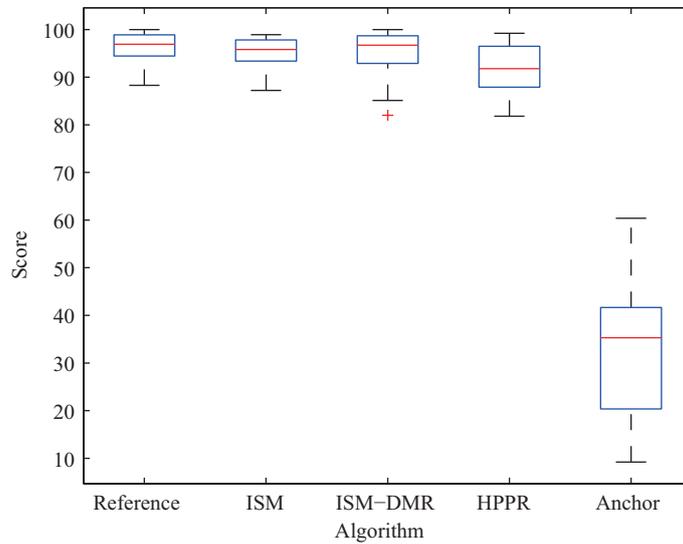


图 5 (网络版彩图) 各个算法 (ISM, ISM-DRM 和 HPPR) 在各个场景和激励下的平均测听结果

Figure 5 (Color online) Listening test results averaged across all simulated scenarios and different stimuli for the traditional algorithms (ISM and ISM-DRM) and the proposed HPPR approach

ISM-DRM 混响模拟方法, 首先产生出相应的 RIR, 然后与“干”信号进行卷积操作, 从而产生混响声音信号. 在 HPPR 混响模拟中, “干”信号与 ISM 建模的早期冲激响应卷积产生早期混响信号, “干”信号直接输入反馈延时网络来产生后期混响, 最后将两者相加模拟出最终混响信号. 在主观测听实验中, 共选用 3 种“干”信号: 女性声音、歌曲和钢琴声. 10 名听力正常的被试 (7 男 3 女) 参加了主观测听实验. 在实验过程中, 每个被试需要听 135 个激励信号 (9 房间  $\times$  3 信号  $\times$  5 算法), 这些激励信号被分为 27 组, 每组包含 5 个信号 (3 种算法模拟得到的混响信号、锚 (anchor) 信号和参考信号). 参考信号是由 ISM 建模的 RIR 与“干”信号卷积得到, 其他 4 种待测信号为: 隐藏的参考信号、锚信号、ISM-DRM 建模 RIR 与“干”信号卷积得到的混响信号、HPPR 模拟的混响信号. 其中, 锚信号是参考信号通过 ITU-RBS.1534 标准定义的低通滤波器得到的 [20]. 利用个人电脑, 实验选取在安静的房间内进行, 测试过程中, 所有激励信号通过 TDH-39 耳机以随机顺序播放给被试进行主观测听. 整个测试过程按照 ITU-RBS.1534 标准进行, 实验所用 Matlab 评测页面及听者受试环境如图 4 所示.

表 1 执行时间加速因子  $R_{CPU}$   
 Table 1 Execution time speed-up factor  $R_{CPU}$

Parameter	Value						
$T_{60}$ (s)	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$R_{CPU}$	1.15	1.34	1.54	1.49	1.40	1.23	1.36

在所有模拟房间和激励源下的主观测听结果表示于图 5. 从图 5 可以看出: 参考混响信号明显不同于锚信号, 根据 ITU-RBS.1534 标准, 此主观测听实验是有效的; 用 3 种待测混响模拟方法产生得到的混响信号都具有很好的主观混响感知效果, 达到了优良水平<sup>[20]</sup>. 为检验信号类型, 房间场景和混响算法对测听结果的影响, 对测听实验结果进行 3 因素方差分析 (ANOVA). 由于方差分析旨在找出 3 种方法产生混响的感知区别, 因此在 ANOVA 分析的过程中, 锚信号和参考信号未考虑在内. ANOVA 分析结果表明: 不同混响模拟方法不存在显著差异 [ $F(2, 18) = 0.47$ ]; 即从统计意义上来讲, 被试对 3 种混响模拟方法产生的混响信号没有感受到明显差异. 因此, HPPR 混响模拟方法与传统的基于卷积运算的混响模拟方法具有相同的听觉感知效果.

### 3.3 计算速度分析

为了对 HPPR 混响模拟方法的计算速度进行分析, 我们对 140 个房间混响进行了模拟, 其中混响时间  $T_{60}$  从 0.1 s 变化到 0.7 s 分布, 间隔为 0.1 s; 每个混响时间条件下, 对 20 个房间混响进行了模拟. 每个房间的体积在  $[30, 100]$  m<sup>3</sup> 随机选择, 声源和接收位置在房间内随机分布. 对于任意给定的  $T_{60}$ , 分别记录 ISM-DRM 和 HPPR 混响模拟方法在 20 个不同房间中混响模拟所需的 CPU 时间. 由于相对于标准 ISM 方法, ISM-DRM 方法计算量更小, 速度更快, 因此在计算速度对比分析中没有考虑 ISM 方法. 在分析中, ISM-DRM 方法和 HPPR 混响模拟方法中, 混合时间都定义为  $\Delta_c = 15$  dB.

在对计算速度进行分析时, 我们采用了执行时间加速因子  $R_{CPU}$  分析, 其定义为 ISM-DRM 产生混响所需时间与 HPPR 产生混响所需时间之比. 在给定的混响时间  $T_{60}$  条件下, 对模拟产生的 20 个不同房间混响所需平均时间的加速因子列于表 1. 表 1 中的结果表明: 相对于 ISM-DRM 混响模拟方法, HPPR 混响模拟方法的计算速度更快, 计算速度大约提高了 30%. HPPR 混响模拟方法计算量的减小主要得益于: (1) 早期混响模拟中, ISM 仅需模拟产生很小阶数的 RIR; (2) 后期混响由“干”信号通过 FDN 直接产生, 不再需要非常耗时的卷积运算.

## 4 结论

本论文提出了一种基于物理特征和感知特征的 HPPR 混合混响模拟方法. HPPR 混响模拟方法中, 早期混响通过卷积 ISM 建模的短早期 RIR 与“干”信号模拟产生; 后期混响利用 FDN 模拟产生. FDN 的参数由早期混响的 EDR 参数化预测模型自动计算确定, 确保了早期混响和后期混响在时—频域的平滑过渡. 数值模拟实验和主观测听实验都验证了此方法的有效性, 相对于目前最佳的 ISM-DRM 混响模拟方法, HPPR 混响模拟方法具有相同的混响感知效果, 但具有更快的计算速度和更大的模拟灵活性.

## 参考文献

- 1 Valimaki V, Parker J D, Savioja L, et al. Fifty years of artificial reverberation. *IEEE Trans Audio Speech*, 2012, 20: 1421–1448
- 2 Svensson P, Kristiansen U R. Computational modelling and simulation of acoustic spaces. In: *Proceedings of AES 22nd Conference on Virtual, Synthetic, and Entertainment Audio*, Espoo, 2002. 11–30
- 3 Funkhouser T A, Carlbom I, Elko G P G, et al. Interactive acoustic modeling for complex environments. *J Acoust Soc Am*, 1999, 105: 1357
- 4 Schroeder M R. Natural-sounding artificial reverberation. *J Audio Eng Soc*, 1962, 10: 219–223
- 5 Schroeder M R. Digital simulation of sound transmission in reverberant spaces. *J Acoust Soc Am*, 1970, 47: 424–431
- 6 Moorer J A. About this reverberation business. *Comput Music J*, 1979, 3: 13–28
- 7 Stautner J, Puckette M. Multi-channel reverbrators. *Comput Music J*, 1982, 6: 52–65
- 8 Jot J M, Chaigne A. Digital delay networks for designing artificial reverberators. In: *Proceedings of Audio Engineering Society Convention 90*, Paris, 1991. 19–22
- 9 Rindel J H. Computer simulation techniques for acoustical design of rooms. *Acoust Aust*, 1995, 2: 81–86
- 10 Murphy D, Stewart R. A hybrid artificial reverberation algorithm. In: *Proceedings of Audio Engineering Society Convention 122*, Vienna, 2007
- 11 Lehmann E, Johansson A. Diffuse reverberation model for efficient image-source simulation of room impulse responses. *IEEE Trans Audio Speech*, 2010, 18: 1429–1439
- 12 Primavera A, Cecchi S, Romoli L, et al. An advanced implementation of a digital artificial reverberator. In: *Proceedings of Audio Engineering Society Convention 130*, London, 2011
- 13 Primavera A, Cecchi S, Piazza F, et al. An efficient time varying hybrid reverberator for room acoustic simulation. In: *Proceedings of the 2nd IEEE China Summit and International Conference on Signal and Information Processing*, Xi'an, 2014. 217–221
- 14 Allen J, Berkley D. Image method for efficiently simulating small-room acoustics. *J Acoust Soc Am*, 1979, 65: 943–950
- 15 Bernfeld B. Absorption of sound by air: a personal calculator program. *J Audio Eng Soc*, 1980, 28: 521–524
- 16 Rocchesso D, Smith J O. Circulant and elliptic feedback delay networks for artificial reverberation. *IEEE Trans Audio Speech*, 1997, 5: 51–63
- 17 Jot J M. An analysis/synthesis approach to real-time artificial reverberation. In: *Proceedings of IEEE Conference on Acoustics, Speech, and Signal Processing*, San Francisco, 1992. 221–224
- 18 Lehmann E, Johansson A. Prediction of energy decay in room impulse responses simulated with an image-source model. *J Acoust Soc Am*, 2008, 124: 269–277
- 19 Sabine W C, Egan M D. *Collected papers on acoustics*. *J Acoust Soc Am*, 1994, 95: 3679–3680
- 20 Stoll G, Kozamernik F. A method for subjective listening tests of intermediate audio quality. ITU-R BS-1534, 2001

## A hybrid physically- and perceptually-based approach for reverberation simulation

XU HuaXing, XIA RiSheng, LI JunFeng\* & YAN YongHong

*Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China*

\*E-mail: lijunfeng@hcl.ioa.ac.cn

**Abstract** Reverberation simulation in many areas including live music and computer games has important applications. Traditional physical and perceptual-based reverberation simulation techniques require a high computational cost and lack the specific characteristics of the room, respectively. This paper presents a hybrid artificial reverberation synthesis method (HPPR): the early reflections are generated by convolving with the impulse response modeled by the image source method (ISM) based on the physical characteristics; the late reverberation is produced by the feedback delay network (FDN) based on the perceptual characteristics. Furthermore, we pro-

pose a method to estimate the energy decay relief (EDR) of the modeled early reflection and thus compute the parameters of the feedback delay network using the method. This process results in the smooth transition from the early reflections to the late reverberation in the time-frequency domain. Numerical simulation and subjective hearing experiments confirm the effectiveness of the proposed method.

**Keywords** reverberation simulation, image-source method, feedback delay network, energy decay relief, room acoustics



**XU HuaXing** was born in 1988. He received a B.E. degree from JiLin University in 2012. Since September 2012, he has pursued his doctoral degree at The Key Laboratory of Speech Acoustics and Content Understanding, Chinese Academy of Sciences, Beijing. His research interests are primarily 3D audio processing.



**XIA RiSheng** was born in 1978. He has worked as an embedded system designer for several years, concentrating on speech communication products. Currently he is a researcher at the Key Lab of Speech Acoustics and Content Understanding, Institute of Acoustics, Chinese Academy of Sciences. His research interests include 3D audio processing, speech coding, and embedded system design.



**LI JunFeng** graduated from Xidian University in 2003 and received a Ph.D. degree in Information Science from the Japan Advanced Institute of Science and Technology (JAIST) in March 2006. From April 2006, he was a post-doctoral research fellow at the Research Institute of Electrical Communication (RIEC), Tohoku University. From April 2007 to July 2010, he was an assistant professor at the School of Information Science, JAIST. Since August 2010, he has been a professor at the Institute of Acoustics, Chinese Academy of Sciences. His research interests include psychoacoustics, speech signal processing, and 3D audio technology. Dr. Li is now serving as the Subject Editor for Speech Communication and the Editor for IEICE Trans. on Fundamentals of Electronics, Communication and Computer Sciences.



**YAN YongHong** received a B.E. degree from the Electronic Engineering Department of Tsinghua University in 1990 and a Ph.D. degree in Computer Science and Engineering from Oregon Graduate Institute (OGI) of Science and Engineering in 1995. From 1995 to 1998, he worked at OGI as an assistant professor, Associate Director, and associate professor of the Center for Spoken Language Understanding. From 1998 to 2001 he worked as the Principal Engineer for Intel Microprocessors Research Lab and Director and Chief Scientist of the Intel China Research Center. In 2002 he returned to China to work for the Chinese Academy of Sciences. He is a professor and Director of the Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics, Chinese Academy of Sciences. His research interests include large vocabulary speech recognition, speaker/language recognition, and audio signal processing.