

# 双扬声器双耳回放系统串扰消除算法的改进

朱天一 卢晶 陈锴

(南京大学声学研究所近代声学教育部重点实验室, 南京, 210093)

**摘要:** 双扬声器系统的一个典型应用就是实现局部范围的双耳声回放, 其目标是在听者每只耳朵处回放相应通道的声信号, 需要解决的最重要问题是消除两个扬声器发出信号的串扰。为达到此目的, 需设计串扰消除滤波器。通用的串扰消除滤波器是通过声源到人耳处的传递函数矩阵求逆得到的, 其面临的显著问题是会导致较为严重的频谱染色。频谱染色会带来诸多负面效应比如系统鲁棒性变差、最佳听音区范围的缩小和动态范围损失等。本文重点研究基于实测传递函数的两种串扰消除优化方案, 通过实验在频谱染色和串扰消除方面对两种方法的性能进行对比, 并进一步通过双耳时间差的计算分析判断回放系统的有效性。

**关键词:** 串扰消除; 频谱染色; 规整参数

**中图分类号:** O422      **文献标志码:** A

## Improvement of Crosstalk Cancellation for Stereo Reproduction System Based on Two Loudspeakers

Zhu Tianyi, Lu Jing, Chen Kai

(MOE Key Lab of Modern Acoustics, Institution of Acoustics, Nanjing University, Nanjing, 210093, China)

**Abstract:** One of the most important functions of the two-loudspeaker system is to realize stereo playback in a specific region. Through sending the ipsilateral channel of stereo signals to the corresponding ear of the listener. The fundamental challenge is to cancel the crosstalk of signals from two loudspeakers. Usually, crosstalk cancellation filter is employed based on the inverse of the transfer matrix from sources to ears. Nevertheless, common used crosstalk cancellation will cause severe spectral coloration which brings many negative effects, e. g. weakening system robustness, shrinking sweet spot area and losing considerable dynamic range. Therefore, two different regularization methods based on measured transfer functions are investigated. Experiments are carried out to compare the integrated performance of the spectral coloration and the crosstalk cancellation performance. Interaural time difference (ITD) is also utilized to validate the efficacy of the improved stereo sound reproduction system.

**Key words:** crosstalk cancellation; spectral coloration; regularization parameter

## 引 言

基于双扬声器的双耳声回放的最终目标是在双耳处实现声像重建<sup>[1]</sup>, 也就是在听者的双耳耳道入

口处产生只包含双耳声信号对应通道的声信息,使得听者准确获得双耳声录音中的双耳信息,进而产生逼真的虚拟声场图像<sup>[2]</sup>。然而在通过双扬声器进行回放的过程中,由于每个扬声器产生的声场同时能被同侧和对侧的耳朵接收,相应的串扰不可避免,为此必须有效地实现串扰消除。

传统的双耳声回放通过耳机实现。此方法相对简单,因为它避免了扬声器回放中出现的串扰问题,但是在具体的回放效果上,耳机回放也存在诸多缺陷:(1)头中效应(双耳感受到的虚拟声源位置位于头内部)<sup>[3]</sup>; (2)由于传递函数的不匹配、听者头部移动、缺乏骨传导以及换能器引起的耳道共振等因素的影响,听者无法感受到真实的三维声场图像<sup>[4,5]</sup>。与耳机回放相比,扬声器回放技术存在潜在的优势:(1)有效避免了头中效应,因为声源位于距离听者一定距离的位置;(2)回放时的骨传导过程以及听者躯干、耳道的衍射效应增强了声场回放的真实性。

扬声器回放系统的早期研究可追溯到 20 世纪 70 年代<sup>[6-8]</sup>,随着对串扰消除机制的深入研究,文献[9]提出了基于系统传递函数矩阵求逆的串扰消除方法。但是串扰消除技术本身存在困难限制了其大规模推广应用:(1)对于头部运动的敏感和有限的最佳听音区<sup>[10,12]</sup>; (2)对室内声反射的敏感<sup>[13]</sup>; (3)扬声器阵列设置变动带来的影响<sup>[14,15]</sup>。此外,串扰消除方法的另一大制约因素是频谱染色<sup>[16]</sup>。频谱染色现象会带来许多问题:(1)在频谱增益峰值位置对应的是传递函数求逆问题的病态情况<sup>[17]</sup>,在病态情况下串扰消除效果对误差的敏感度很大,当处于最佳听音区内的听者头部位置有微调时就会造成双耳接收信号的较大失真,即使在听者试听位置理想不动的情况下,最佳听音区外的听者也会听到明显的增益过大现象;(2)频谱染色的高增益对声源设备造成大的负担;(3)造成声信号动态范围的较大损失。文献[16]提出了两种优化的串扰消除方案,通过增加一个规整因子减弱频谱染色的影响,但该文的讨论基于理想传递函数的分析,缺乏对实际系统的分析讨论和实验验证。本文以实测传递函数为基础,通过仿真和实验对比两种优化方案——恒定参数优化方案和与频率相关参数优化方案在频谱染色和串扰消除方面的效果。为进一步分析优化结果的有效性,还在试验结果中引入双耳时间差(Interaural time difference, ITD)来检验双耳回放系统的听觉特性。

## 1 双扬声器回放系统数学模型

如图 1 所示,信号从扬声器发出到被人耳接收的过程可表示为

$$\mathbf{P} = \mathbf{CHd} = \mathbf{Cv} \quad (1)$$

式中:  $\mathbf{P} = \begin{bmatrix} P_L(j\omega) \\ P_R(j\omega) \end{bmatrix}$  为人双耳接收到的声压向量;  $\mathbf{C} = \begin{bmatrix} C_{LL}(j\omega) & C_{LR}(j\omega) \\ C_{RL}(j\omega) & C_{RR}(j\omega) \end{bmatrix}$  为两个扬声器到双耳的系统传递函数频响矩阵;  $\mathbf{H} = \begin{bmatrix} H_{LL}(j\omega) & H_{LR}(j\omega) \\ H_{RL}(j\omega) & H_{RR}(j\omega) \end{bmatrix}$  为串扰消除滤波器的频响矩阵,  $\mathbf{d} = \begin{bmatrix} D_L(j\omega) \\ D_R(j\omega) \end{bmatrix}$  为待回放的信号频谱,  $\mathbf{v} = \begin{bmatrix} V_L(j\omega) \\ V_R(j\omega) \end{bmatrix}$  为经过串扰消除滤波器滤波后馈给扬声器的信号频谱。

$$\mathbf{v} = \mathbf{Hd} \quad (2)$$

为了使人耳接收到的声压信号和待回放的信号一致,须满足

$$\mathbf{p} = \mathbf{CHd} = \mathbf{d} \quad (3)$$

$$\mathbf{CH} = \mathbf{I} \quad (4)$$

$$\mathbf{H} = \mathbf{C}^{-1} \quad (5)$$

所以,串扰消除滤波器的设计核心就是对系统传递函数频响矩阵的

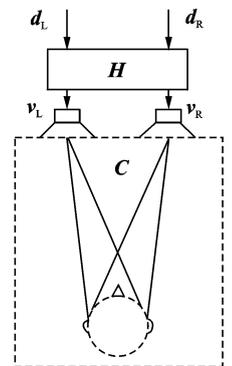


图 1 双扬声器回放系统模型  
Fig. 1 Schematic diagram of stereo reproduction system based on two loudspeakers

求逆。此外,为了评价回放系统的串扰消除效果,引入矩阵  $\mathbf{R}$  并定义了左右耳串扰消除量

$$\mathbf{R} = \mathbf{C}\mathbf{H} \quad (6)$$

$$\chi_L(\omega) = 20\lg\left(\frac{|R_{LL}(j\omega)|}{|R_{LR}(j\omega)|}\right) \quad (7)$$

$$\chi_R(\omega) = 20\lg\left(\frac{|R_{RR}(j\omega)|}{|R_{RL}(j\omega)|}\right) \quad (8)$$

式中: $\chi_L(\omega)$ ,  $\chi_R(\omega)$  分别为左右耳串扰消除量,  $R_{LL}(j\omega)$ ,  $R_{LR}(j\omega)$ ,  $R_{RL}(j\omega)$ ,  $R_{RR}(j\omega)$  为式(1)中  $\mathbf{R}$  矩阵的 4 个元素。在理想情况下根据式(4),  $\mathbf{R}$  应该为单位矩阵,此时左右耳串扰消除量均趋于无穷大。在左右耳传递函数完全对称的情况下  $\chi_L(\omega)$ ,  $\chi_R(\omega)$  应该相等,但在实际实验系统中,传递函数的对称性很难保证。

## 2 频谱染色问题以及衡量指标

从两个声源传播到耳朵的声波在空间中会形成干涉现象。在某些频率,人耳处的声压与声源强度发生弱耦合,也就是说声源激励不能有效地转化为声压。为了使得声压达到合适幅度,必须提高此频率下声源的激励增益,而这种增益提高在通常情况下能达到 30 dB,随之带来的就是严重的频谱染色现象<sup>[16]</sup>。这主要表现为馈给扬声器的信号频谱在某些频率存在峰值。

由式(2),馈给扬声器的信号由待回放信号经串扰消除滤波器滤波后得到。馈给扬声器信号的频谱能量可表示为

$$P_v(\omega) = \|\mathbf{v}\|_2 = \|\mathbf{H}\mathbf{d}\|_2 \quad (9)$$

由线性代数中范数相容性知识易知

$$\|\mathbf{H}\mathbf{d}\|_2 \leq \|\mathbf{H}\|_2 \|\mathbf{d}\|_2 \quad (10)$$

式中: $\|\mathbf{d}\|_2$  为待回放信号的能量。由式(10)可知,当待回放信号能量一定时,馈给扬声器信号能量取决于  $\|\mathbf{H}\|_2$ ,所以本文用  $\|\mathbf{H}\|_2$  作为衡量频谱染色的指标。因此将频谱染色定义为

$$S(\omega) = \|\mathbf{H}\|_2 \quad (11)$$

式中: $S(\omega)$  为不同频率处的频谱染色强弱。

## 3 频谱染色问题的优化方案

### (1) 恒定参数优化方案

综合考虑回放误差和频谱染色现象,定义代价函数

$$J = \mathbf{e}^H \mathbf{e} + \beta \mathbf{v}^H \mathbf{v} \quad (12)$$

式中: $\mathbf{e} = \mathbf{P} - \mathbf{C}\mathbf{H}\mathbf{d}$  为回放误差; $\beta$  为优化因子,反映了代价函数中考虑频谱染色的权重因子。

恒定参数优化方案的出发点是对式(12)中代价函数  $J$  的最小化。利用最小二乘法得到对应代价函数最小的最优解

$$\mathbf{H} = (\mathbf{C}^H \mathbf{C} + \beta \mathbf{I})^{-1} \mathbf{C}^H \quad (13)$$

式中: $\mathbf{I}$  为单位矩阵。利用式(13)对串扰消除滤波器采用恒定参数优化后的频谱染色  $S(\omega)$  能被抑制在  $\frac{1}{2\sqrt{\beta}}$  以下。

### (2) 与频率相关参数优化方案

与频率相关参数优化方案同样是利用式(13)得到串扰消除滤波器,与恒定参数优化方案的区别在于优化因子  $\beta$  随着频率而改变,变化的原则是设定一个阈值  $\Gamma$ ,当不同频率下的  $S(\omega)$  不超过  $\Gamma$  时, $\beta$  取 0 即不加优化;当  $S(\omega)$  超过  $\Gamma$  时,则取适当的  $\beta$  使得  $S(\omega) = \Gamma$ 。

## 4 实验仿真结果和性能比较

为了验证串扰消除系统的性能和两种优化方案对频谱染色的抑制作用,本文在普通房间中分别利用两种优化方案进行了扬声器阵列回放实验,实验装置示意图如图 2 所示。待回放信号经过串扰消除滤波器滤波后,经过功率放大器放大反馈给扬声器。回放后的信号通过 B&K 4128 型人工头在配套的 B&K 公司的 Pulse 平台上完成采集。

串扰消除滤波器的性能好坏取决于系统传递函数的准确程度,在本实验中传递函数的测量是将输入扬声器的信号与人工头采集信号进行维纳滤波匹配得到。下面分别就恒定参数优化方案和与频率相关参数优化方案的串扰消除效果和频谱染色情况给出实验结果。

### 4.1 频谱染色情况比较

图 3 显示了两种优化方案处理后的频谱染色情况的比较,为较为清楚地展示对比效果,优化参数经过选择,保证频谱染色的峰值不超过 7 dB。从图 3 中可以看出,两种方案均能很明显地抑制频谱染色,相对而言,恒定参数优化方案频谱染色的峰谷波动更为明显,而与频率相关参数优化方案则能将超过阈值的频率部分的频谱染色强制“削平”,其余频率保持不变。更进一步的分析表明,两种优化方案的频谱染色情况均随着优化力度增加而得到抑制。

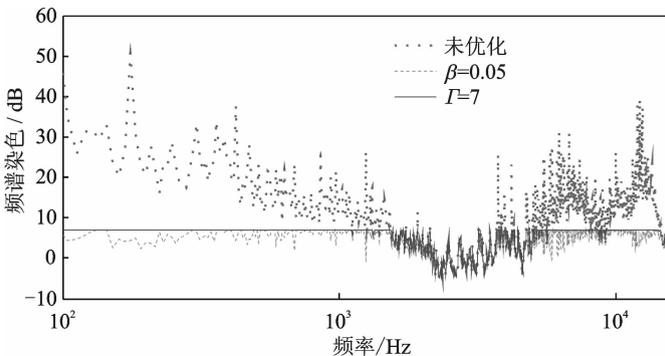


图 3 两种参数优化方案下的频谱染色对比

Fig. 3 Comparison of spectral coloration using two regularization methods

### 4.2 串扰消除效果比较

由于实际回放系统左右无法完全对称,左右耳串扰消除量不会完全相等,在比较两种优化方案的串扰消除效果时本文只考虑左耳串扰消除量。在这里我们采用了两种方法对串扰效果进行评价。第 1 种方法是直接利用式(7,8)得到左耳串扰消除量  $\chi_L(\omega)$ ,这也是理论上能够得到的串扰消除效果;第 2 种方法是分别将相同功率的左、右单通道白噪声信号进行回放后,用两次回放后信号左通道的功率谱之差来表征串扰消除效果。相比第 1 种方法,第 2 种方法更真实反映了实际系统的串扰消除效果。

由图 4 可以看出,理论上不论采取哪种优化方案,随着  $\beta$  的减小或  $\Gamma$  的增大,串扰消除效果都会明显下降,而且与频率相关的参数优化方案的串扰消除效果优于恒定参数优化方案,因为明显可以看出与频率相关参数优化在更大的频率范围内达到了较好的串扰消除( $\geq 40$  dB)。然而图 5 的结果则表示,在

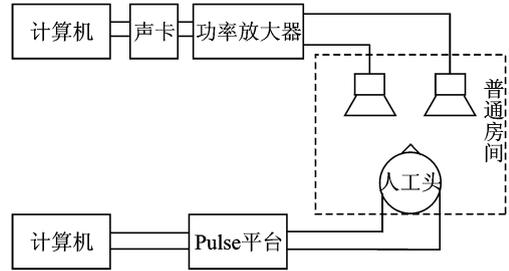


图 2 双扬声器回放实验系统框图

Fig. 2 Experiment system of stereo reproduction system based on two loudspeakers

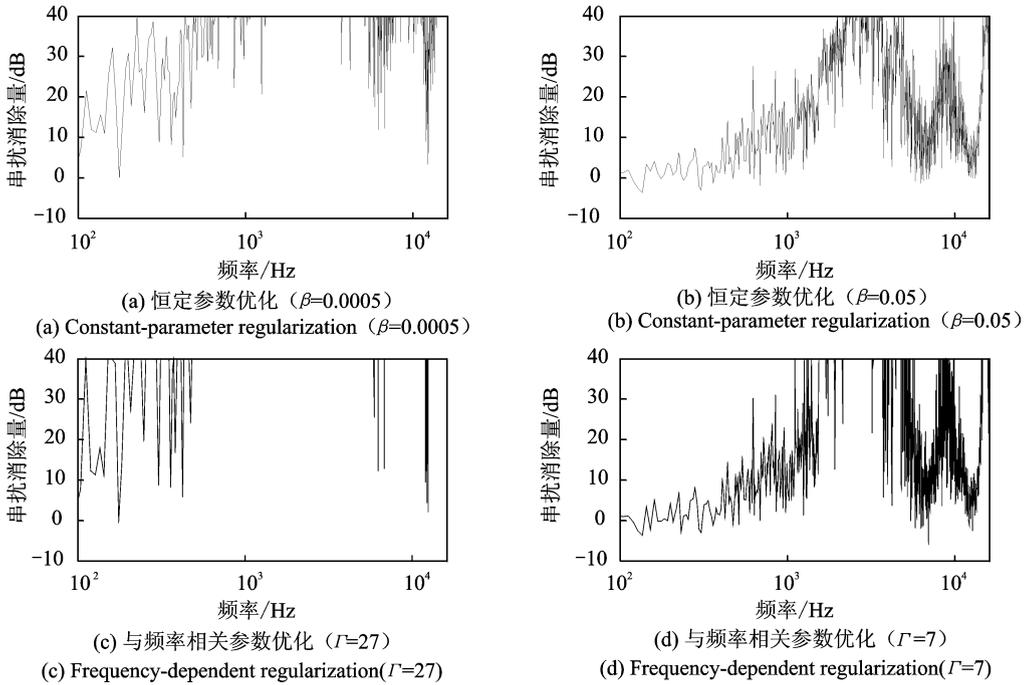


图 4 两种优化方式在不同优化参数选择下的串扰消除理论值

Fig. 4 Theoretical performance of crosstalk cancellation with different parameters

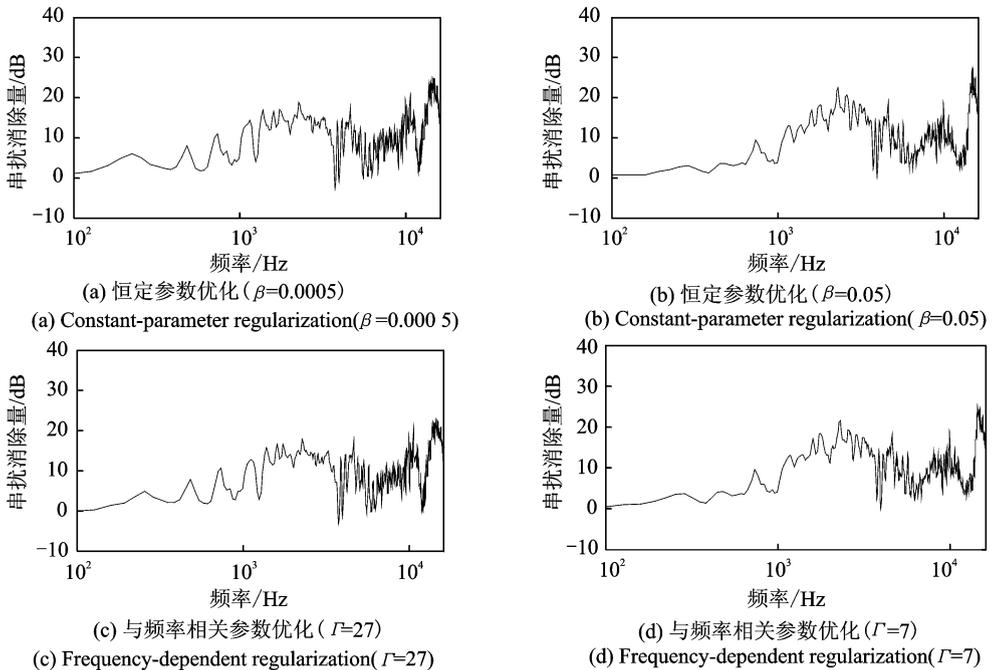


图 5 两种优化方式在不同优化参数选择下的串扰消除实测值

Fig. 5 Measured performance of crosstalk cancellation with different parameters

实际回放过程中两种优化方案整体的串扰消除效果大幅度下降,而且随着优化参数的改变,串扰消除性能并没有明显的削弱,仅仅是串扰消除效果极差的频率范围有所增加,反映在图中就是串扰消除曲线的“谷”变宽,而这些串扰消除极差的频率对应的正是系统传递函数求逆的病态特征频率。

理论上,随着频谱染色效应的改善,串扰消除效果变差合理。由式(5,13)表明,优化的过程实际上就是为了抑制频谱染色而人为在传递函数频响矩阵求逆时引入误差,所以串扰消除效果的变差在预期之中。而在一些频率,对应系统传递函数求逆的病态情况,串扰消除效果对串扰消除滤波器的误差非常敏感,所以表现在图中形成了一系列“谷”。然而实测值却不符合以上推论,即使在较为保守的优化参数选择条件下(比如  $\beta=0.0005$ ),串扰消除量也没有达到理想的效果。造成以上情况的原因主要是实际系统的不确定性,即式(1)测量得到的系统传递函数与实际传递函数仍然存在一定程度的失配,失配的原因,一方面在于测量时维纳滤波的匹配程度不够好;另一方面是测量过程中的室内环境和系统位置扰动带来的实际传递函数变化。

### 4.3 双耳时间差比较

串扰消除的最终目的是实现对待回放信号的高保真回放。理论上串扰消除效果越好,对待回放信号的还原度越高,回放后的双耳信息也越接近原信号。ITD是对声源进行空间定位的重要信息之一,实验中将包含不同ITD信息的原信号分别采用两种优化方案回放。比较不同优化方案和优化参数选择下回放信号与原信号的ITD差异,结果见表1,2。

表1,2中的原信号是指希望在双耳处重建的声信号,其他的分别为采用不同优化参数的优化方案回放后的信号。从表1,2数据可以发现,两种优化方案在等效优化参数选择下对ITD的还原准确度没有明显差异。不过,当优化参数过于激进( $\beta$ 增加到0.05,阈值 $\Gamma$ 下降到7后)时,ITD发生明显偏离,意

表1 恒定参数优化方案下ITD比较

Tab. 1 Comparison of ITD using constant-parameter regularization method

原信号序号	原信号	$\beta=0.0005$	$\beta=0.005$	$\beta=0.05$	$\beta=0.158$
1	-0.259 4	-0.244 1	-0.228 9	-0.167 8	-0.091 6
2	-0.625 6	-0.625 6	-0.640 9	0.061 0	0.045 8
3	-0.686 6	-0.686 6	-0.701 9	0.061 0	0.030 5

表2 与频率相关参数优化方案下ITD比较

Tab. 2 Comparison of ITD using frequency-dependent regularization method

原信号序号	原信号	$\Gamma=26$	$\Gamma=17$	$\Gamma=7$	$\Gamma=2$
1	-0.259 4	-0.244 1	-0.244 1	-0.167 8	-0.091 6
2	-0.625 6	-0.610 4	-0.625 6	-0.671 4	-0.045 8
3	-0.686 6	-0.686 6	-0.686 6	0.061 0	0.045 8

味着回放效果变差。这与之前的实测串扰消除效果吻合。ITD对声源的定位作用体现在低频段大约1.5 kHz以下。随着优化力度增加,低频段的串扰消除效果变差,直接导致ITD的偏差变大。

## 5 结束语

本文重点从实验上比较了不同优化方案对双扬声器回放系统的频谱染色的抑制作用和对回放效果的影响。研究表明,恒定参数优化方案和与频率相关参数优化方案对于频谱染色的抑制效果均十分明显,在串扰消除效果上也不存在明显的差异,但需要选取合适的优化参数,在频谱染色抑制和系统

回放效果之间取得平衡。优化参数过于保守则频谱染色不能得到有效抑制,优化力度过于激进则会明显削弱回放效果。考虑到恒定参数优化方案的算法复杂度明显低于与频率相关参数优化方案,因此在工程应用上是一个相对简单而有效的优化方案。

### 参考文献:

- [1] 胡瑞敏, 王晓晨, 张茂盛, 等. 三维音频技术综述 [J]. 数据采集与处理, 2014, 29(5): 661-676.  
Hu Ruimin, Wang Xiaochen, Zhang Maosheng, et al. Review on three-dimension audio technology [J]. Journal of Data Acquisition and Processing, 2014, 29(5): 661-676.
- [2] Hugonnet C, Walder P A. Stereophonic sound recording [M]. Chichester, UK: John Wiley & Sons, Inc, 1998.
- [3] Blauert J. Spatial hearing [M]. Cambridge, US: The MIT Press, 1983.
- [4] Takeuchi T, Nelson P, Hamada H. Robustness to head misalignment of virtual sound imaging systems [J]. Journal of the Acoustical Society of America, 2001, 109(3): 958-971.
- [5] Akeroyd M A, Chambers J, Bullock D, et al. The binaural performance of a cross-talk cancellation system with matched or mismatched setup and playback acoustics [J]. Journal of the Acoustical Society of America, 2007, 121(2): 1056-1069.
- [6] Bauer B B. Stereophonic earphones and binaural loudspeakers [J]. Journal of the Audio Engineering Society, 1961, 9(2): 148-151.
- [7] Atal B S, Hill M, Schroeder M R. Apparent sound source translator [P]. US: 3236949, 1966-2-22.
- [8] Damaske P. Head-related two-channel stereophony with loudspeakers reproduction [J]. Journal of the Acoustical Society of America, 1971, 50(4): 1109-1115.
- [9] Nelson P A, Hamada H, Elliott S J. Adaptive inverse filters for stereophonic sound reproduction [J]. IEEE Transactions on Signal Processing, 1992, 40(7): 1621-1632.
- [10] Ward D B, Elko G. Effect of loudspeaker position on the robustness of acoustic crosstalk cancellation [J]. IEEE Signal Processing Letters, 1999, 6(5): 106-108.
- [11] Rose J, Nelson P, Rafaely B, et al. Sweet spot size of virtual acoustic imaging systems at asymmetric listener locations [J]. Journal of the Acoustical Society of America, 2002, 112(5): 1992-2002.
- [12] Bai M R, Lee C C. Objective and subjective analysis of effects of listening angle on crosstalk cancellation in spatial sound reproduction [J]. Journal of the Acoustical Society of America, 2006, 120(4): 1976-1989.
- [13] Ward D B. On the performance of acoustic crosstalk cancellation in a reverberant environment [J]. Journal of the Acoustical Society of America, 2001, 110(2): 1195-1198.
- [14] Kirekby O, Nelson P A, Hamada H. Local sound field reproduction using two closely spaced loudspeakers [J]. Journal of the Acoustical Society of America, 1998, 104(4): 1973-1981.
- [15] Kirekby O, Nelson P A, Hamada H. The 'stereo dipole': A virtual source imaging system using two closely spaced loudspeakers [J]. Journal of the Audio Engineering Society, 1998, 46(5): 387-395.
- [16] Choueiri E Y. Optimal crosstalk cancellation for binaural audio with two loudspeakers [EB/OL]. <http://www.princeton.edu/3D3A/Papers.html>, 2015-1-13/2015-1-27.
- [17] Nelson P A, Rose J F W. Errors in two-point sound reproduction [J]. Journal of the Acoustical Society of America, 2005, 118(1): 193-204.

### 作者简介:



朱天一(1991-),男,硕士研究生,研究方向:音频信号处理,E-mail:805938932@qq.com。



卢晶(1977-),男,副教授,研究方向:音频信号处理、声场计算。



陈锴(1980-),男,博士,研究方向:声信号处理。

