

◇ 研究报告 ◇

# 两扬声器虚拟声重放的双耳声压控制与定位性能\*

李枚婧 谢菠菘<sup>†</sup> 刘路路

(华南理工大学物理与光电学院 声学研究所 广州 510641)

**摘要:** 在两扬声器虚拟声重放中,通过精确重构双耳声压而产生不同的空间听觉感知。其重放的定位性能应该是由双耳声压控制的代价和稳定性所共同决定的。过去研究主要对双耳声压控制的稳定性进行分析,并以此作为扬声器布置和信号处理的依据。该文研究表明仅对双耳声压的稳定性分析是不足以完全衡量扬声器虚拟声重放的定位性能的。进一步采用虚拟声信号处理滤波器响应平均功率对双耳声压控制的代价进行分析。结果表明,缩窄左右对称扬声器布置的张角或采用非对称扬声器布置会明显增加产生侧向目标虚拟源时的双耳声压控制代价。虚拟源(虚拟声像)定位实验表明,双耳声压控制代价增加会引起虚拟源定位缺陷。实际应用中,为了有效产生侧向虚拟源,应避免采用过窄张角(如立体声偶极)和非对称的扬声器布置。

**关键词:** 虚拟声; 定位; 双耳声压控制

中图法分类号: O428; O429 文献标识码: A 文章编号: 1000-310X(2023)01-0001-11

DOI: 10.11684/j.issn.1000-310X.2023.01.001

## Binaural sound pressure control and localization performance in virtual sound reproduction using two loudspeakers

LI Meiqiang XIE Bosun LIU Lulu

(Acoustic Lab, School of Physics and Optoelectronics, South China University of Technology, Guangzhou 510641, China)

**Abstract:** In virtual sound reproduction via two loudspeakers, binaural pressures are controlled by two loudspeakers to create various spatial auditory perceptions. Localization performance in reproduction is determined by the stability and effort of binaural pressure control. Stability in binaural pressure control is analyzed in previous studies to provide the basis for the design of loudspeaker configuration and signal processing. It is indicated in present work that analysis on the stability only is not enough to evaluate the localization performance in reproduction completely. The effort of binaural pressure control is further analyzed using the average power of filters of virtual sound signal processing. Results indicate that symmetric loudspeaker configurations with narrow span angle and asymmetric loudspeaker configuration increase the effort of binaural pressure control for lateral virtual source. Virtual source localization experiment indicates that an increase of the effort of binaural pressure control leads to localization defect. In practice, to recreate lateral virtual source in reproduction effectively, symmetric loudspeaker configurations with narrow span angle (such as stereo dipole) and asymmetric loudspeaker configuration should be avoided.

**Keywords:** Virtual sound; Localization; Binaural pressure control

2021-11-30 收稿; 2022-03-07 定稿

\*国家重点研发计划项目(2018YFB1403800)

作者简介: 李枚婧(1997-),女,广西玉林人,硕士研究生,研究方向:电声与信号处理。

<sup>†</sup>通信作者 E-mail: phbsxie@scut.edu.cn

## 0 引言

虚拟声或虚拟听觉重放是一种空间声重放技术<sup>[1]</sup>。基于头相关传输函数(Head related transfer function, HRTF)信号处理,虚拟声期望在耳机或扬声器重放中重构目标声源产生的双耳声压,从而产生期望的听觉感知。其中,与传统的多通路声比较,采用扬声器的虚拟声重放的硬件结构相对简单,采用一对前方的扬声器即有可能在中心倾听位置产生前半水平面的虚拟源与听觉事件<sup>[2]</sup>。因而扬声器虚拟声重放广泛地应用于不适宜布置多个扬声器的重放场合,如电视机、计算机及汽车内声重放,特别是用于将多通路环绕声信号转换为两扬声器或少量扬声器重放<sup>[3-4]</sup>。

在物理上,采用扬声器的虚拟声重放等价于用若干扬声器(最少两个)对双耳声压进行控制,使其尽可能接近目标声压<sup>[5-6]</sup>。因而最终重放性能应该是由双耳声压控制的代价和稳定性所共同决定的。代价是指扬声器是否能准确且高效地控制双耳声压,而稳定性是指双耳声压的控制对物理条件(如倾听位置)小的改变是否稳定。

过去对扬声器虚拟声重放稳定性已有许多研究工作,并以此作为重放设计和评价的重要手段。分析稳定性的常用方法是计算扬声器到双耳声学传输矩阵的条件数<sup>[7-8]</sup>。普遍结果表明,稳定性和实际的扬声器布置、频率有关。过去的工作对不同扬声器布置的虚拟声重放稳定性进行了分析,提出了各种改善稳定性的信号处理和扬声器布置方法,并取得一定的成效。Kirkeby等<sup>[9]</sup>提出在信号处理中引入与频率有关的正则化系数来改善稳定性。Kirkeby等<sup>[10]</sup>还提出采用张角为 $10^\circ$ 的一对扬声器(立体声偶极)重放而改善中频稳定性的方法。而Takeuchi等<sup>[11]</sup>提出采用3对不同张角的扬声器进行分频段重放(最佳源分布)而改善重放稳定性的方法。另外,也有采用功率均衡抵消零点而改善稳定性和音色的方法<sup>[2,3,12]</sup>。

另一方面,双耳声压控制代价确定了扬声器虚拟声重放能否产生期望的听觉事件。但过去对这方面的研究工作较少。最近有工作对扬声器到双耳的声学传输矩阵的独立模态进行分析,并分析了双耳声压模态的控制效率<sup>[13]</sup>,但并未将分析结果和实际的虚拟源(虚拟声像)定位效果相联系。目前也未见

有关不同扬声器布置双耳声压控制代价及其对重放虚拟源定位影响的报道。

另外,过去许多工作主要针对扬声器相对倾听者对称布置的情况。但在一些实际应用中,扬声器相对倾听者是左右非对称布置的<sup>[14]</sup>。这种非对称的布置可能会影响最终的重放效果。虽然过去也有工作研究了对称扬声器布置重放时,倾听偏离中心倾听位置微小距离而引起的听声区域大小问题<sup>[15]</sup>;但专门涉及非对称扬声布置的虚拟声重放详细分析较少,特别未见有针对非对称扬声器布置所产生的双耳声压控制与虚拟源定位缺陷问题的研究,相应的虚拟源定位实验的工作也较少。

本文以两扬声器虚拟声重放为例,探讨了不同张角对称扬声器布置和非对称扬声器布置的声压控制的代价和稳定性问题,并通过主观心理声学实验研究其与虚拟源定位缺陷之间的关系。

## 1 两扬声器虚拟声重放的原理

由于前半水平面扬声器布置的虚拟声重放只能产生前半水平面的虚拟源<sup>[2]</sup>,本文约定采用相对于倾听者头中心的水平面极坐标系统。倾听者的头中心位于坐标系统的原点,水平面任一点的位置用坐标 $(r, \theta)$ 表示。其中, $r$ 为水平点到原点的距离; $-180^\circ < \theta \leq 180^\circ$ 为水平方位角, $\theta = -90^\circ$ 、 $\theta = 0^\circ$ 和 $\theta = 90^\circ$ 分别表示正左、正前和正右方。

这里采用双耳声压控制的方法导出两扬声器虚拟重放的信号<sup>[5-6]</sup>。如果目标声源和扬声器都位于 $r > 1.0$  m的远场距离,则它们到双耳的HRTF近似和距离无关。事实上,当目标声源和扬声器到倾听者的距离大于0.5 m时,它们到双耳的HRTF随距离变化已比较慢<sup>[16]</sup>,因而可以略去表示距离的变量 $r$ 。如图1所示,两扬声器分别布置在倾听者前方 $\theta_1$ 与 $\theta_2$ 的方位角。假定目标虚拟源和重放扬声器都位于远场距离,两扬声器到双耳的4个声学传输函数分别为 $H_{L1} = H_{L1}(\theta_1, f)$ 、 $H_{R1} = H_{R1}(\theta_1, f)$ 、 $H_{L2} = H_{L2}(\theta_2, f)$ 和 $H_{R2} = H_{R2}(\theta_2, f)$ ,其中 $f$ 是频率。再假定扬声器1、扬声器2的信号分别为 $E_1 = E_1(f)$ 和 $E_2 = E_2(f)$ ,则两扬声器产生的双耳声压可用以下的矩阵公式表示:

$$\begin{bmatrix} P'_L \\ P'_R \end{bmatrix} = \begin{bmatrix} H_{L1} & H_{L2} \\ H_{R1} & H_{R2} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}. \quad (1)$$

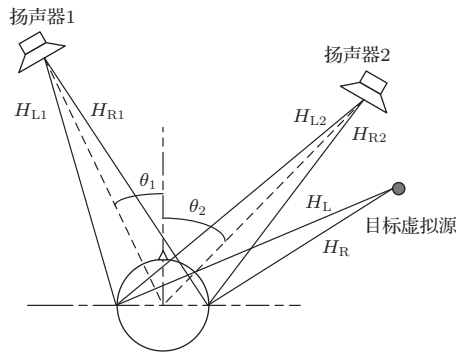


图1 两扬声器虚拟重放原理图

Fig. 1 Schematic of virtual sound reproduction via two loudspeakers

另一方面,水平面方位角为 $\theta_S$ 的目标声(虚拟)源到左耳和右耳的HRTF分别用 $H_L = H_L(\theta_S, f)$ 和 $H_R = H_R(\theta_S, f)$ 表示。它在双耳产生的目标声压可用以下公式表示:

$$\begin{bmatrix} P_L \\ P_R \end{bmatrix} = \begin{bmatrix} H_L \\ H_R \end{bmatrix} E_0(f), \quad (2)$$

其中, $E_0(f)$ 表示源强度或正比于源强度的信号。如果令式(1)左边的重放双耳声压与式(2)给出的双耳声压相等,即可求解出两扬声器的重放信号:

$$\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} H_L \\ H_R \end{bmatrix} E_0(f), \quad (3)$$

其中,串声消除矩阵 $\mathbf{A}$ 为两扬声器到双耳声学传输矩阵 $\mathbf{H}$ 的逆:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} H_{L1} & H_{L2} \\ H_{R1} & H_{R2} \end{bmatrix}^{-1} = \mathbf{H}^{-1}. \quad (4)$$

或者:

$$\mathbf{A} = \frac{1}{H_{L1}H_{R2} - H_{L2}H_{R1}} \begin{bmatrix} H_{R2} & -H_{L2} \\ -H_{R1} & H_{L1} \end{bmatrix}. \quad (5)$$

式(3)的右边的信号处理包括两步。第一步是用一对目标虚拟源方向的HRTF对信号 $E_0(f)$ 滤波,而 $E_0(f)$ 正比于目标声源强度,从而合成期望的双耳声信号。第二步是经串声消除矩阵 $\mathbf{A}$ 处理,得到两扬声器重放信号,且重放时可以得到目标声源一致的双耳声压。

双耳声信号合成与串声消除处理可以合并为一步,称为听觉传输(transaural)合成处理<sup>[5]</sup>。将

式(4)和式(5)代入式(3),两扬声器的重放信号也可以写为

$$E_1 = G_L(\theta_S, f)E_0(f), \quad E_2 = G_R(\theta_S, f)E_0(f), \quad (6)$$

其中:

$$G_L(\theta_S, f) = \frac{H_{R2}H_L(\theta_S, f) - H_{L2}H_R(\theta_S, f)}{H_{L1}H_{R2} - H_{L2}H_{R1}},$$

$$G_R(\theta_S, f) = \frac{-H_{R1}H_L(\theta_S, f) + H_{L1}H_R(\theta_S, f)}{H_{L1}H_{R2} - H_{L2}H_{R1}}. \quad (7)$$

式(6)表明,将信号 $E_0(f)$ 用一对听觉传输滤波器 $G_L = G_L(\theta_S, f)$ 和 $G_R = G_R(\theta_S, f)$ 滤波后,即可得到两扬声器的重放信号。以上就是两扬声器虚拟或听觉传输重放的基本原理。

由式(4)可以看出,上述信号处理只有在两扬声器到双耳声学传输矩阵 $\mathbf{H}$ 是可逆的条件下才是有效的。当矩阵 $\mathbf{H}$ 是病态而接近不可逆时,串声消除和两扬声器虚拟重放将变得不稳定。虽然这种情况下可以用各种正则化方法近似求矩阵的逆<sup>[17]</sup>,但正则化方法本身就引入了误差,影响双耳声压控制的精确性。

## 2 双耳声压控制稳定性与代价分析

### 2.1 分析方法

扬声器虚拟声重放的性能应该同时由双耳声压控制的稳定性与代价进行评价。双耳声压控制稳定性可看成系统存在一定干扰情况下的稳定性(鲁棒性)问题。好的稳定性体现在当虚拟重放过程存在一定干扰或误差时,例如倾听者头部的移动、头部的转动以及实际HRTF与信号处理用的HRTF不匹配,虚拟重放的效果不易受干扰所影响。由式(3)和式(4),串声消除矩阵 $\mathbf{A}$ 是扬声器到双耳的传输矩阵 $\mathbf{H}$ 的逆。因而扬声器虚拟声重放对微小扰动的稳定性主要是与传输矩阵 $\mathbf{H}$ 逆的稳定性有关。一般情况下,可以用传输矩阵的条件数来衡量虚拟重放在干扰作用下的稳定性<sup>[1]</sup>,它定义为

$$\kappa(\mathbf{H}) = \sigma_1/\sigma_2, \quad (8)$$

其中, $\sigma_1$ 和 $\sigma_2$ 分别表示传输矩阵 $\mathbf{H}$ 的最大奇异值和最小奇异值。在理想条件下, $\kappa(\mathbf{H}) = 1$ 。 $\kappa(\mathbf{H})$ 越小(越接近1),说明虚拟重放在干扰下的稳定性越高;反之 $\kappa(\mathbf{H})$ 越大,说明虚拟重放的稳定性越差。

对于双耳声压控制的代价,是衡量扬声器是否能准确且高效地控制双耳声压的物理指标,这里提出采用一对听觉传输滤波器在一定频率范围内平均功率进行分析,与“双耳声重放的驱动信号功率大小”密切相关。根据式(6)和式(7),两扬声器虚拟重放的信号是通过一对听觉传输滤波器  $G_L = G_L(\theta_S, f)$  和  $G_R = G_R(\theta_S, f)$  对输入信号  $E_0$  滤波得到,它们重放后理论上可产生期望(与目标声源一致)的双耳声压。滤波器响应  $G_L = G_L(\theta_S, f)$  和  $G_R = G_R(\theta_S, f)$  在一定频率范围的平均功率为(以dB为单位):

$$W(\theta_S) = 10 \times \lg \frac{\left\{ \sum_k [|G_L(\theta_S, f_k)|^2 + |G_R(\theta_S, f_k)|^2] \right\}}{K \Delta f} \text{ (dB)}, \quad (9)$$

其中,对数内对离散频率  $f_k$  求和是在所讨论的频率范围内进行,  $\Delta f$  是离散频率的间隔,  $K$  为求和频率的点数。当式(6)输入信号是在所讨论的频带内具有平坦的单位功率谱  $|E_0|^2 = |E_0(f)|^2 = 1$  (如带通白噪声),上式表示准确控制双耳压所需要的两扬声器信号的平均功率,它与扬声器的布置有关,也和目标虚拟源的方向  $\theta_S$  有关。 $W(\theta_S)$  越小表示控制双耳声压的平均效率越高,代价越小。反之,  $W(\theta_S)$  增加表示准确控制双耳压比较困难,很大一部分的声压振幅在双耳处干涉叠加而抵消,同时也会引起双耳外某些位置的声压振幅叠加增强。在后面的计算中,将0.1~10 kHz和0.1~1.5 kHz两个频率范围对  $W(\theta_S)$  进行计算。前者是由于10 kHz以上的高频成分对应的波长很短,在扬声器重放中很难对双耳声压进行控制,而实际重放中采用一般采用非个性化HRTF处理,进一步地增加了高频误差。并且这些成分对扬声重放的定位贡献不大。后者是由于考虑频率1.5 kHz以下的双耳因素(双耳时间差)是侧向虚拟源定位的主导因素<sup>[18]</sup>。

后面将采用式(8)的条件数和式(9)的滤波器响应平均功率进行分析。对比分析了下面3组扬声器布置的情况:

(1) 第一组,不同张角的对称扬声器布置。包括张角分别为10°、30°和60°的3种布置情况,即扬声器的布置分别为  $\theta_1 = -30^\circ, \theta_2 = 30^\circ$ ;  $\theta_1 = -15^\circ, \theta_2 = 15^\circ$ ;  $\theta_1 = -5^\circ, \theta_2 = 5^\circ$ 。

(2) 第二组,张角固定为60°的非对称扬声器布置。包括两种情况,即  $\theta_1 = -15^\circ, \theta_2 = 45^\circ$ ;  $\theta_1 = 0^\circ, \theta_2 = 60^\circ$ , 并和第一组得到的对称布置  $\theta_1 = -30^\circ, \theta_2 = 30^\circ$  的情况比较。

(3) 第三组,扬声器2固定,扬声器1在不同方位角的非对称扬声器布置。包括两种情况,其中扬声器2的方位角为  $\theta_2 = 30^\circ$ , 扬声器1方位角分别为  $-15^\circ$  和  $0^\circ$  情况,并和第一组得到的对称布置  $\theta_1 = -30^\circ, \theta_2 = 30^\circ$  的情况比较。

需要说明的是,在实际的重放中,对称听声位置窄,多数情况倾听者无法做到一直固定在理想对称听声位置,例如倾听者往左右大幅度移动,这时,左右扬声器相对倾听者的方位角是非对称的;同时,它们相对倾听者的距离也不相同。这种距离差异会导致左右扬声器到倾听者的直达声延时和衰减的差异。但给定左右扬声器与倾听者位置之间的几何关系,直达声延时和衰减的差异是很容易通过适当的扬声器信号延时和增益控制补偿的。为简单起见,同时不失一般性,上面第二组和第三组的分析假定左右扬声器到倾听者的距离相等,但方位角是非对称的。

为了考虑由头部对声波的散射效应,仿真采用了边界元计算的得到的KEMAR人工头(配DB-060/061小耳廓和颈部,但不包括躯干)的远场HRTF数据<sup>[19]</sup>。HRTF采样频率为44.1 kHz,长度为512点,声源距离为1.5 m。当实际声源距离大于0.5 m时,该HRTF数据可作为合理的近似。

## 2.2 双耳声压控制稳定性的分析结果

图2(a)给出了第一组扬声器布置情况下,式(8)计算得到的传输矩阵  $\mathbf{H}$  的条件数的结果。分析中只给出了频率范围为0.1~10 kHz的仿真结果。从图2(a)中可以看到,对称扬声器布置的情况下,在1.3 kHz以下的低频带,条件数随扬声器布置的张角减小而增加。特别是张角为10°(立体声偶极子)的扬声器布置,在低频带的条件数明显比另外两种布置的条件数大,因而低频双耳声压控制的稳定性较差。在中高频带,不同张角布置的条件数都出现了波动。特别是在3.5~7.0 kHz的频段,张角为10°的扬声器布置条件数最小,表现出相对好的双耳声压控制稳定性。而张角为30°扬声器布置是考虑2.5 kHz以下频段条件数的一种折中。这与

过去的分析以及采用不同张角的扬声器进行分频段重放(最佳源分布)而改善重放稳定性的研究相符合<sup>[11]</sup>。

图2(b)给出了第二组布置情况下,传输矩阵  $\mathbf{H}$  的条件数的结果。可以看到固定张角的情况下,3种扬声器布置的条件数和稳定性是类似的,但随着扬声器布置的非对称性增加(例如  $\theta_1 = 0^\circ$ 、 $\theta_2 = 60^\circ$  的情况),条件数略有增加,稳定略为变差。

图2(c)给出了第三组布置情况下,传输矩阵  $\mathbf{H}$  的条件数的结果。可以看到固定扬声器2方向的

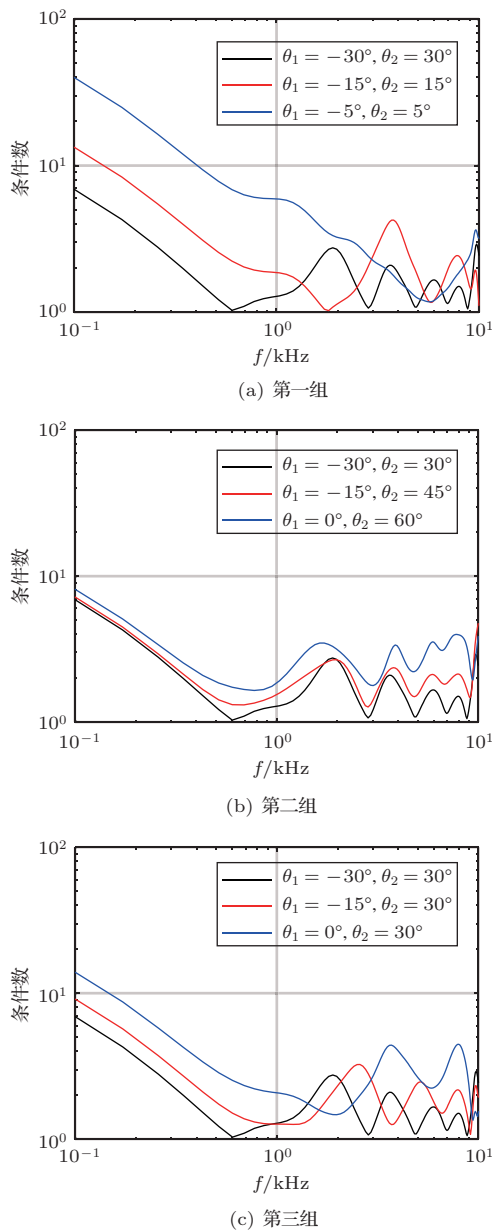


图2 3组扬声器布置条件数的分析结果

Fig. 2 Analysis result of condition number of three groups of loudspeaker configuration

情况下,在1.3 kHz以下,  $\theta_1 = -30^\circ$ 、 $\theta_2 = 30^\circ$  对称扬声器布置的条件数最小;  $\theta_1 = -15^\circ$ 、 $\theta_2 = 30^\circ$  的扬声器布置次之;  $\theta_1 = 0^\circ$ 、 $\theta_2 = 30^\circ$  的扬声器布置最大。在1.3 kHz以上的频率范围,条件数随频率涨落,但平均的规律和1.3 kHz低频类似。条件数增加从而稳定性变差也主要是两扬声器之间的张角缩窄所引起,非对称性的影响较少。

### 2.3 双耳声压控制代价的分析

图3给出了0.1~10 kHz频率范围的滤波器响应平均功率。图3(a)是第一组扬声器布置的结果。可以看出,对于同一扬声器布置,随目标虚拟源偏离正前方  $\theta_S = 0^\circ$  而趋向侧向  $\theta_S = \pm 90^\circ$ ,需要的平均功率增加。 $\theta_1 = -5^\circ$ 、 $\theta_2 = 5^\circ$  的扬声器布置尤为明显。并且随着扬声器之间张角的缩窄,平均功率明显增加。例如,对于  $\theta_S = 90^\circ$  的虚拟源,  $\theta_1 = -5^\circ$ 、 $\theta_2 = 5^\circ$  和  $\theta_1 = -30^\circ$ 、 $\theta_2 = 30^\circ$  扬声器布置所需要的功率分别是 -9.6 dB 和 -15.5 dB,两者相差了 5.9 dB。因而,在两扬声器对称布置的情况下,控制双耳声压而模拟侧向虚拟源较模拟前方虚拟源困难。缩窄两扬声器之间的张角明显增加侧向虚拟源双耳声压控制的代价。

图3(b)是第二组扬声器布置的结果。可以看出,当两扬声器布置之间的张角固定时,与对称布置比较,非对称的布置对同侧虚拟源的滤波器响应平均功率影响较小,但明显增加了布置异侧虚拟源的滤波器响应平均功率。例如,对于  $\theta_1 = 0^\circ$ 、 $\theta_2 = 60^\circ$  非对称布置,其布置同侧 ( $\theta_S = 90^\circ$ ) 虚拟源的滤波器响应平均功率与  $\theta_1 = -30^\circ$ 、 $\theta_2 = 30^\circ$  对称布置类似,布置异侧 ( $\theta_S = -90^\circ$ ) 虚拟源的滤波器响应平均功率却从对称布置的 -15.5 dB 增加到 -10.6 dB,增幅为 4.9 dB。因而非对称布置主要是影响异侧虚拟源双耳声压的控制的代价。

图3(c)是第三组扬声器布置的结果。这是上面第一组和第二组结果的综合。即缩窄两扬声器之间的张角明显增加侧向虚拟源双耳声压控制的代价,非对称布置主要是影响异侧虚拟源双耳声压的控制代价。例如,对于  $\theta_1 = 0^\circ$ 、 $\theta_2 = 30^\circ$  非对称布置,其布置同侧 ( $\theta_S = 90^\circ$ ) 虚拟源的滤波器响应平均功率与  $\theta_1 = -30^\circ$ 、 $\theta_2 = 30^\circ$  对称布置类似,布置异侧 ( $\theta_S = -90^\circ$ ) 虚拟源的滤波器响应平均功率却从对称布置的 -15.5 dB 增加到 -10.4 dB,增幅为 5.1 dB。

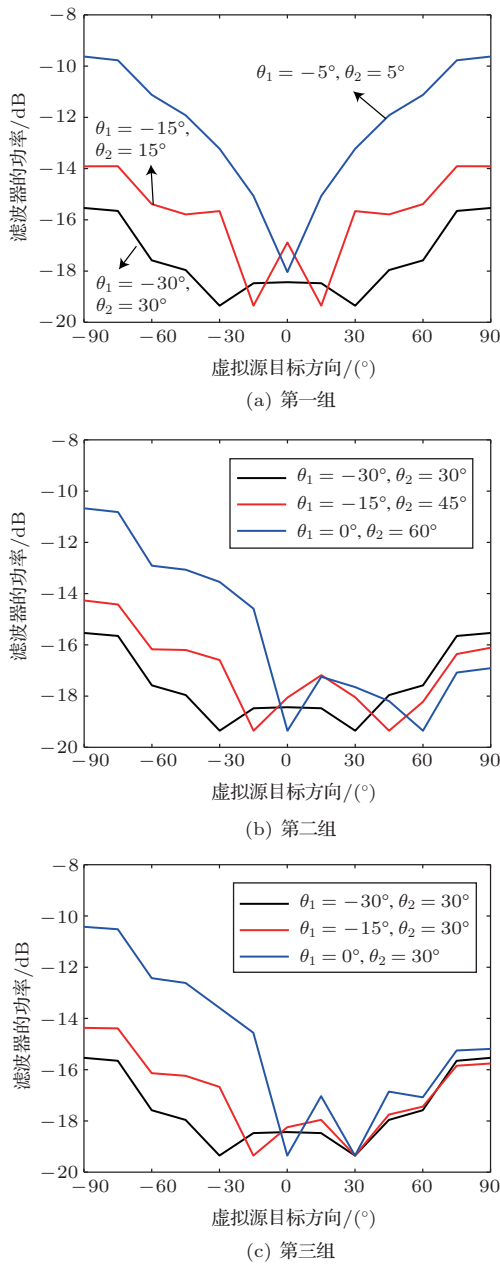


图3 0.1~10 kHz 频率范围, 3组扬声器布置的滤波器响应平均功率

Fig. 3 Filter response average power of three groups of loudspeaker configuration in the frequency range of 0.1-10 kHz

由于1.5 Hz以下的低频双耳时间差(Interaural time difference, ITD)是扬声器重放中虚拟源侧向定位的主导因素,进一步对0.1~1.5 kHz频率范围内3组扬声器布置的滤波器响应平均功率进行分析,结果如图4所示。图4(a)是第一组扬声器布置的结果。与前面有类似结果,对于同一扬声器布置,随着目标虚拟源偏离正前方 $\theta_S = 0^\circ$ 而趋向侧向 $\theta_S = \pm 90^\circ$ ,需要的平均功率增加。对于 $\theta_1 = -5^\circ$ 、

$\theta_2 = 5^\circ$ 的扬声器布置,这种趋势更为明显,同时相比0.1~10 kHz频率范围,0.1~1.5 kHz频率范围内需要的平均功率更高。例如,对于 $\theta_S = 90^\circ$ 的虚拟源, $\theta_1 = -5^\circ$ 、 $\theta_2 = 5^\circ$ 和 $\theta_1 = -30^\circ$ 、 $\theta_2 = 30^\circ$ 扬声器布置所需要的功率分别是-4.2 dB和-16.2 dB,两者相差了12 dB。

图4(b)是第二组扬声器布置的结果。可以看出,与对称布置比较,非对称的布置对同侧虚拟源的

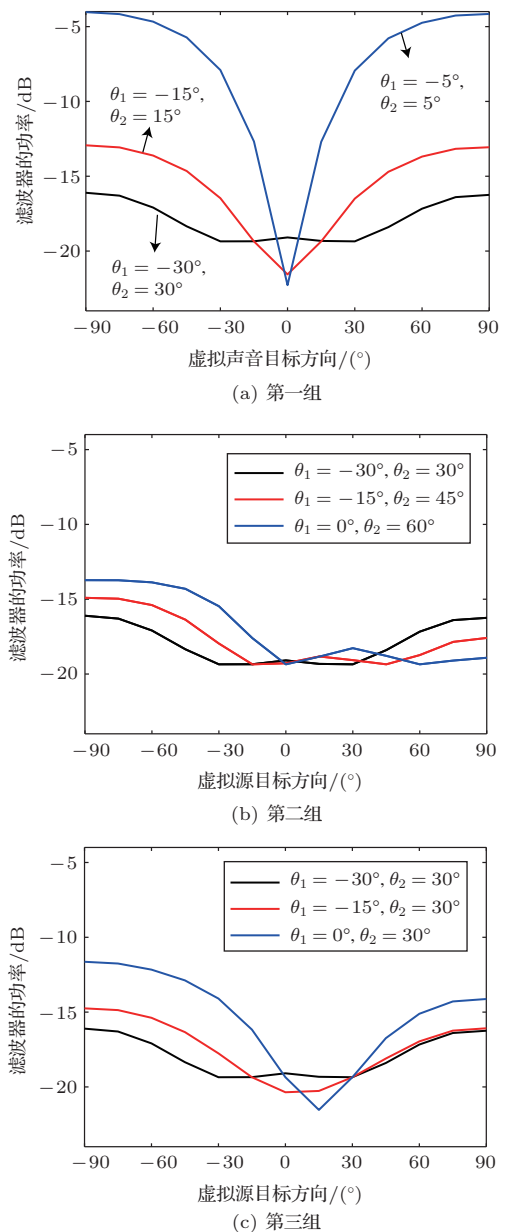


图4 0.1~1.5 kHz 频率范围, 3组扬声器布置的滤波器响应平均功率

Fig. 4 Filter response average power of three groups of loudspeaker configuration in the frequency range of 0.1-1.5 kHz

滤波器响应平均功率影响较小, 仅在  $75^\circ$ 、 $90^\circ$  有一定影响, 但明显增加了布置异侧虚拟源的滤波器响应平均功率。例如, 对于  $\theta_1 = 0^\circ$ 、 $\theta_2 = 60^\circ$  非对称布置, 异侧 ( $\theta_S = -90^\circ$ ) 虚拟源的滤波器响应平均功率从对称布置的  $-16.2$  dB 增加到  $-13.9$  dB, 增幅为  $2.3$  dB。因而非对称布置主要是影响异侧虚拟源双耳声压的控制的代价。

图4(c)是第三组扬声器布置的结果。缩窄两扬声器之间的张角明显增加侧向虚拟源双耳声压控制的代价, 非对称布置主要是影响异侧虚拟源双耳声压的控制代价。例如, 对于  $\theta_1 = 0^\circ$ 、 $\theta_2 = 30^\circ$  非对称布置, 布置异侧 ( $\theta_S = -90^\circ$ ) 虚拟源的滤波器响应平均功率却从对称布置的  $-16.2$  dB 增加到  $-11.7$  dB, 增幅为  $4.5$  dB。

### 3 虚拟源定位实验

#### 3.1 实验设计

前面的分析表明, 两扬声器虚拟声重放的实际定位性能是与双耳声压控制的代价有关。虚拟源定位实验的目的是证实上述结论。

实验信号包括全频带粉红噪声、 $1.5$  kHz 低通滤波粉红噪声和  $700$  Hz 低通滤波粉红噪声, 信号长度为  $10$  s, 采样频率为  $44.1$  kHz。对前面2.1节所述的3组扬声器布置进行实验。其中对全频带粉红噪声, 对3组中全部扬声器布置进行了实验。而为了简化实验, 对  $1.5$  kHz 和  $700$  Hz 低通滤波粉红噪声, 只对第一组的  $\theta_1 = -30^\circ$ 、 $\theta_2 = 30^\circ$  和  $\theta_1 = -5^\circ$ 、 $\theta_2 = 5^\circ$  扬声器布置、第二组的  $\theta_1 = 0^\circ$ 、 $\theta_2 = 60^\circ$  扬声器布置和第三组的  $\theta_1 = 0^\circ$ 、 $\theta_2 = 30^\circ$  扬声器布置进行了实验。将单通路信号经过式(7)的滤波处理生成  $\theta_S = 0^\circ$ 、 $\pm 15^\circ$ 、 $\pm 30^\circ$ 、 $\pm 45^\circ$ 、 $\pm 60^\circ$ 、 $\pm 75^\circ$ 、 $\pm 90^\circ$  共13个目标方向信号, 再馈给扬声器重放。

实验在混响时间为  $0.15$  s 的听声室内进行。本底噪声不大于  $30$  dB(A)。实验使用 PhilipsMCI500H 扬声器, 所有扬声器距离受试者头中心  $1.45$  m。共8名受试者参加实验, 年龄介于  $22\sim 28$  岁, 且听力正常。实验开始前对受试者进行适当的听声训练。实验中, 鼓励受试者头尽量不动, 判断合成虚拟源方向, 以口头报告的形式记录。每种信号都会随机播放3次, 因此每种信号共有  $3$  次  $\times 8$  人 =  $24$  个判断。最后对结果进行统一的分析处理。

#### 3.2 实验结果与分析

数据反映受试者并未出现前后混乱错误。对每个目标角度的数据进行独立样本 t 检验, 在显著水平  $\alpha = 0.05$  的条件下检验实验结果。

图5为全频带粉红噪声信号的虚拟源定位实验的结果, 图中给出了感知虚拟源平均方位角和标准差。图5(a)为第一组对称扬声器布置的定位实验结果, 总体上看对称扬声器布置的定位结果几乎是对称的。张角为  $60^\circ$  和  $30^\circ$  的两种扬声器布置, 即  $\theta_1 = -30^\circ$ 、 $\theta_2 = 30^\circ$  和  $\theta_1 = -15^\circ$ 、 $\theta_2 = 15^\circ$  布置的定位结果相近。对  $|\theta_S| \leq 60^\circ$  范围的平均感知虚拟源的方位角与目标方位角接近, 可以得到较为理想的定位效果。对  $75^\circ \leq |\theta_S| \leq 90^\circ$  的侧向目标虚拟源, 平均感知虚拟源方向就在  $\pm 60^\circ$  左右(向前方漂移), 产生虚拟源定位错误。对张角为  $10^\circ$  (立体声偶极,  $\theta_1 = -5^\circ$ 、 $\theta_2 = 5^\circ$ ) 扬声器布置, 其  $|\theta_S| \leq 45^\circ$  范围的平均感知虚拟源的方位角与目标方位角接近, 可以得到较为理想的定位效果。但在侧向  $60 \leq |\theta_S| \leq 90^\circ$  范围, 平均感知虚拟源方向就在  $\pm 45^\circ$  左右(向前方漂移)。因而张角为  $10^\circ$  的扬声器布置更容易产生侧向定位错误。以上产生侧向定位错误的规律是和2.3节对双耳声压控制代价的分析一致的, 即定位错误是和双耳声压控制代价的增加相对应的。另外实验过程发现  $10^\circ$  张角的扬声器布置的定位实验对于听者初始位置与扬声器位置的要求精度更高, 稍微偏移会严重影响定位效果。这和过去对立体声偶极重放的理论分析结论不同。

图5(b)为第二组固定张角 ( $60^\circ$ ) 非对称扬声器布置的定位实验结果, 为了比较, 把相同张角对称扬声器布置的结果也重新画在图中。总体上来看, 在非对称布置同侧区域, 平均感知虚拟源方向与对称布置的情况差别不大。在  $\theta_1 = 0^\circ$ 、 $\theta_2 = 60^\circ$  扬声器布置, 对  $\theta_S = 60^\circ$  和  $\theta_S = 90^\circ$  两个目标方向的虚拟源, 平均感知方向比对称布置时更接近目标方向, 因而虚拟源定位错误略为变小。但在非对称布置的异侧区域, 平均感知虚拟源方向向前方漂移比对称布置时更明显, 特别是目标虚拟源方向  $\theta_S = 90^\circ$  的侧向, 这时对应的标准差相对较大。进一步的统计 t 检验也证实了这观察。这里的侧向定位错误的规律也是和2.3节对双耳声压控制代价的分析一致的, 即定位错误是和双耳声压控制代价的增加相对应的。

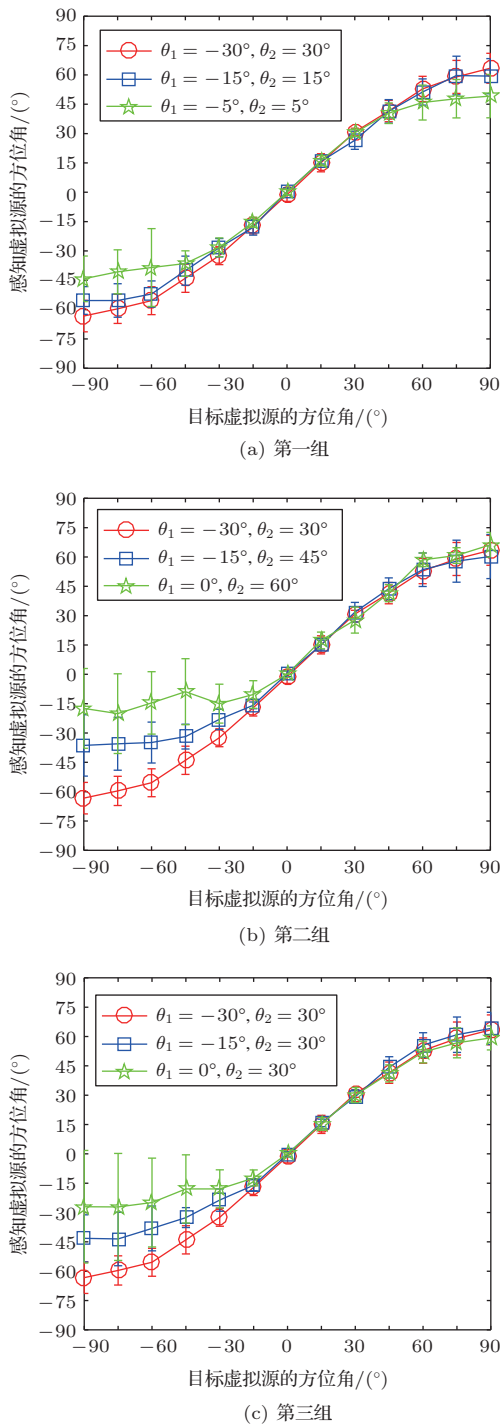


图5 全频带粉红噪声信号的虚拟源定位实验的结果  
Fig. 5 Results of virtual source location experiments for full band pink noise signal

图5(c)为第三组扬声器布置的定位实验结果,这是不同张角非对称布置的综合情况,其中扬声器2的固定在方位角 $\theta_2 = 30^\circ$ ,扬声器1的方位角在 $-15^\circ$ 和 $0^\circ$ 两种情况。为了比较,把对称扬声器布置 $\theta_1 = -30^\circ$ 的结果也重新画在图中。总体上来看,与

第二组的结果类似,非对称布置主要引起异侧的侧向感知虚拟源方向定位错误,感知虚拟源向前漂移,同时稳定性也变差(标准差变大),这应该是缩窄张角带来的影响。这里的侧向定位错误的规律也是和2.3节对双耳声压控制代价的分析一致的,即定位错误是和双耳声压控制代价的增加相对应的。

补充1.5 kHz和700 Hz低通粉红噪声信号的实验是为了探讨不同信号带宽对定位的影响。图6给出了实验的结果。为了比较,把相同扬声器布置的全频带粉红噪声的定位结果也重新画在图中。图6(a)为第一组对称扬声器布置中的 $\theta_1 = -30^\circ$ 、 $\theta_2 = 30^\circ$ 布置的定位实验结果,统计(t-检验)表明全频带粉红噪声与700 Hz低通粉红噪声信号定位结果一致。图6(b)为 $\theta_1 = -5^\circ$ 、 $\theta_2 = 5^\circ$ 对称布置的定位实验结果,统计检验表明1.5 kHz与700 Hz低通粉红噪声信号定位结果无差别,而1.5 kHz、700 Hz低通粉红噪声与全频带粉红噪声的结果差别不大,但在目标虚拟源 $\theta_S = \pm 90^\circ$ 、 $\pm 75^\circ$ 的方向平均感知虚拟源方向的错误略为变小。这是因为高频波长短,容易引起双耳声压控制的误差,如倾听者头部的偏移或者转动都会引起双耳声压发生很大的改变,定位结果不稳定,出现虚拟源漂移现象,所以全频带粉红噪声信号在侧向方向的定位结果偏窄。

图6(c)为 $\theta_1 = 0^\circ$ 、 $\theta_2 = 60^\circ$ 扬声器布置的定位实验结果,统计检验表明1.5 kHz与700 Hz低通粉红噪声信号定位结果差别不大。在非对称布置同侧区域,1.5 kHz、700 Hz低通粉红噪声与全频带粉红噪声的定位结果几乎一致,而在非对称布置异侧区域,平均感知虚拟源方向的错误明显减少,定位效果得到一定的改善。出现这种情况原因有:第一,与高频波长短,容易引起双耳声压控制的误差有关,导致出现虚拟源漂移的现象。第二,从“双耳声压控制的代价”。从 $G_L = G_L(\theta_S, f)$ 和 $G_R = G_R(\theta_S, f)$ 的平均功率角度分析,0.1~10 kHz频带的滤波器功率为 $-10.6$  dB,而0.1~1.5 kHz频带的滤波器功率为 $-13.9$  dB,功率减少了3.2 dB。这也和2.3节对双耳声压控制代价的分析一致的。

图6(d)为 $\theta_1 = 0^\circ$ 、 $\theta_2 = 30^\circ$ 扬声器布置的定位实验结果。定位结果与 $\theta_1 = 0^\circ$ 、 $\theta_2 = 60^\circ$ 扬声器布置趋势一致,1.5 kHz与700 Hz低通粉红噪声信号定位结果差别不大,在非对称布置同侧区域,



1.5 kHz、700 Hz 低通粉红噪声与全频带粉红噪声的定位结果几乎一致，而在非对称布置异侧区域，平均感知虚拟源方向的错误明显减少。出现这种情况原因有：第一，与高频波长短，容易引起双耳声压控制的误差有关，导致出现虚拟源漂移现象。第

二，0.1~10 kHz 频带的滤波器功率为 -10.4 dB，而 0.1~1.5 kHz 频带的滤波器功率为 -11.7 dB，功率减少了 1.3 dB。这里的侧向定位错误的规律也是和 2.3 节对双耳声压控制代价的分析一致的，即定位错误是和双耳声压控制代价的增加相对应的。

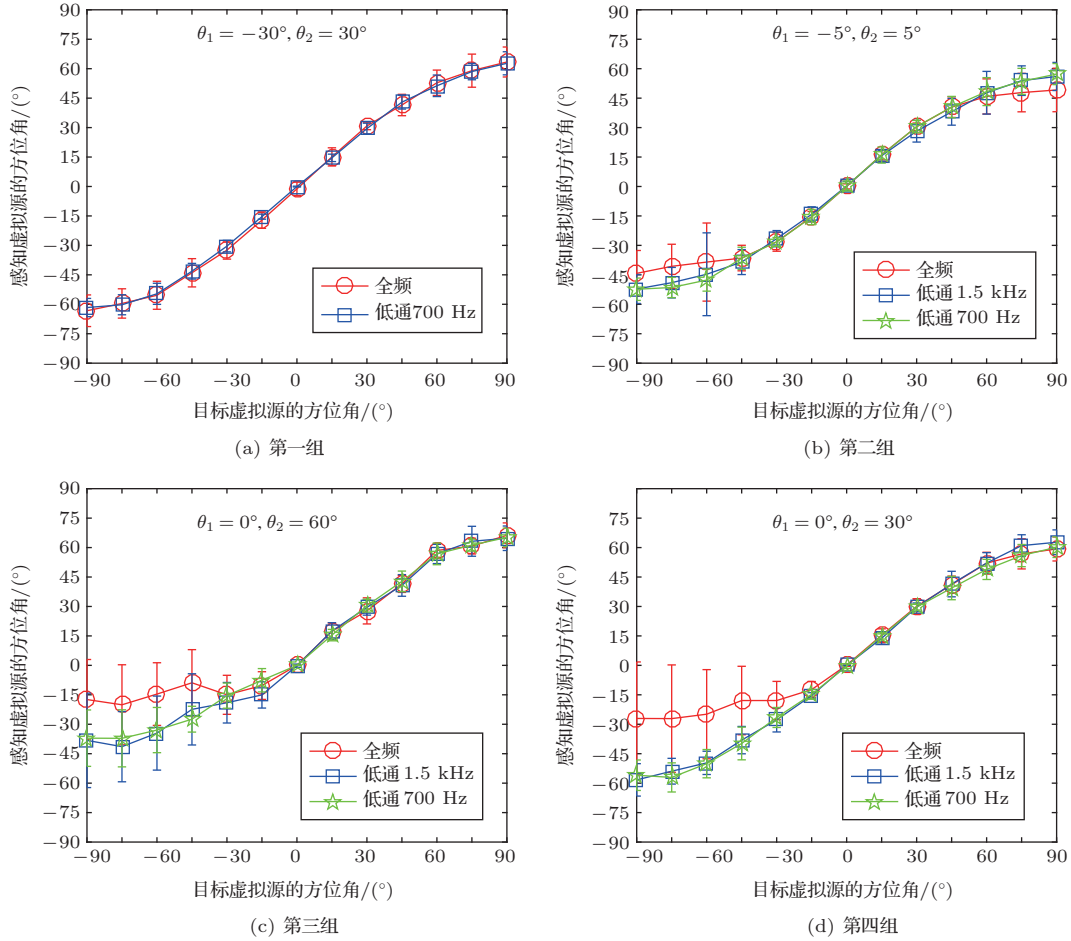


图 6 全频带与低通粉红噪声信号的虚拟源定位实验的结果

Fig. 6 Results of virtual source location experiments for full band and low-pass pink noise signals

#### 4 讨论与结论

两扬声器虚拟重放理论上可以在特定的听声位置产生与目标虚拟源一致的双耳声压，但实际上并不能产生整个水平面的虚拟源。过去的心理声学实验表明，头部转动引起的双耳声压特别是双耳时间差的改变 (1.5 kHz 以下的低频动态因素) 和头部、耳廓等对声波的散射引起的双耳接收声压谱的改变 (5~6 kHz 以上的高频谱因素) 是前后定位的主要因素。对前方两扬声器重放，虽然完美的串声消除理论上能产生期望的双耳声压。但高频声波的波

长很短，且容易受到周围环境影响，并且高频 HRTF 是与个性化有关的。因而在实际中要在 5~6 kHz 以上的实现完美的串声消除是困难的，重放中高频声压 (因为谱因素) 的误差很大。同时前方扬声器重放只能产生与前半水平面声源相匹配的低频动态因素。因而前方扬声器重放只能在前半水平面产生稳定的虚拟源，一般不能在后半水平面产生稳定的虚拟源。除非在非常苛刻的实验条件下，包括个性化 HRTF 进行信号处理，在消声室重放和限制听者头部移动和转动，对于部分倾听者，前方两扬声器能在一定程度上产生整个水平面的虚拟源 [20-21]。在

实际的应用条件下,前方两扬声器重放是不能产生后半平面的虚拟源的。最近的一个实验研究验证了上述的分析与结论<sup>[22]</sup>。但虚拟重放的实际性能是由双耳声压控制的稳定性与代价所决定的,而这种稳定性和代价是和扬声器布置以及频率有关的。过去的研究侧重双耳声压控制的稳定性,并通过扬声器到双耳声学传输矩阵的条件数对其进行分析。过去的分析表明<sup>[23]</sup>,对左右对称的扬声器布置,缩窄两扬声器之间的张角可以改善中频双耳声压控制的稳定性,但却使低频双耳声压控制的稳定性变差。本文的分析进一步表明,在两扬声器张角不变的情况下,非对称的布置对双耳声压控制的稳定性产生一定的影响,使得双耳声压控制的稳定性稍变差。

实际重放的虚拟源定位性能与双耳声压控制的与代价密切相关,并且可以用虚拟声信号处理滤波器在所涉及频率范围内的平均功率进行分析。理论和实验表明,在对称扬声器布置的情况,控制双耳声压而模拟侧向虚拟源较模拟前方虚拟源困难。缩窄两扬声器之间的张角明显增加侧向虚拟源双耳声压控制的代价。而非对称布置主要是影响异侧虚拟源双耳声压的控制的代价,使异侧的感知虚拟源向前偏移。

本文的结果可用于扬声器虚拟声重放的设计。特别是过分窄张角的扬声器布置是不合适产生侧向虚拟源的,因而立体声偶极在产生侧向虚拟源方面有潜在的缺陷。而非对称扬声器布置能产生前方角度范围的虚拟源,但用于产生一定异侧向(异侧)虚拟源就比较困难,在实际应用中应该注意。另外,不同扬声器布置特别是非对称布置对音色的影响以及头部左右移动或者转动后感知性能的改变,今后可继续深入研究。

## 参 考 文 献

- [1] 谢波菽. 头相关传输函数与虚拟听觉[M]. 北京: 国防工业出版社, 2008.
- [2] Liu L L, Xie B S. A high-frequency-band timbre equalization method for transaural reproduction with two frontal loudspeakers[J]. Journal of the Audio Engineering Society, 2022, 70(1/2): 1-14.
- [3] 谢波菽, 师勇, 谢志文, 等. 5.1 通路环绕声的虚拟重放系统[J]. 声学学报, 2005, 30(3): 235-241.
- Xie Bosun, Shi Yong, Xie Zhiwen, et al. Virtual reproducing system for 5.1 channel surround sound[J]. Acta Acustica, 2005, 30(3): 235-241.
- [4] Xie B S, Liu L L, Zhang C Y. Virtual reproduction of surround sound in frontal space using four loudspeakers[J]. Chinese Journal of Acoustics, 2021, 40(2): 155-174.
- [5] Bauck J, Cooper D H. Generalized transaural stereo and applications[J]. Journal of the Audio Engineering Society, 1996, 44(9): 683-705.
- [6] Xie B S. Head-related transfer function and virtual auditory display[M]. 2nd ed. Florida, USA: J. Ross Publishing, 2013.
- [7] Ward D B, Elko G W. Effect of loudspeaker position on the robustness of acoustic crosstalk cancellation[J]. IEEE Signal Processing Letters, 1999, 6(5): 106-108.
- [8] Nelson P A, Rose J F W. Errors in two-point sound reproduction[J]. The Journal of the Acoustical Society of America, 2005, 118(1): 193-204.
- [9] Kirkeby O, Rubak P, Johansen L G, et al. Implementation of crosstalk cancellation networks using warped FIR filters[C]. Audio Engineering Society 16th International conference, 1999, Rovaniemi, Finland, Paper: 16-31.
- [10] Kirkeby 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.
- [11] Takeuchi, Philip A, Nelson P A. Optimal source distribution for binaural synthesis over loudspeakers[J]. The Journal of the Acoustical Society of America, 2002, 112(6): 2786-2797.
- [12] 何璞, 谢波菽, 饶丹. 虚拟声音色均衡信号处理方法的主客观分析[J]. 应用声学, 2006, 25(1): 4-12.
- He Pu, Xie Bosun, Rao Dan. Subjective and objective analyses of timber equalized algorithms for virtual sound reproduction by loudspeakers[J]. Journal of Applied Acoustics, 2006, 25(1): 4-12.
- [13] Hamdan E C, Fazi F M. A modal analysis of multichannel crosstalk cancellation systems and their relationship to amplitude panning[J]. Journal of Sound and Vibration, 2021, 490(2021): 115743.
- [14] Rumsey F. Automotive audio: they know where you sit[J]. Journal of the Audio Engineering Society, 2016, 64(9): 705-708.
- [15] Rose J, Nelson P A, Rafaely B, et al. Sweet spot size of virtual acoustic imaging systems at asymmetric listener locations[J]. The Journal of the Acoustical Society of America, 2002, 112(5): 1992-2002.
- [16] Yu G Z, Wu R X, Liu Y, et al. Near-field head-related transfer-function measurement and database of human subjects[J]. The Journal of the Acoustical Society of America, 2018, 143(3): EL194-198.
- [17] Kirkeby O, Nelson P A. Digital filter design for inversion problems in sound reproduction[J]. Journal of the Audio Engineering Society, 1999, 47(7/8): 583-595.
- [18] Wightman F L, Kistler D J. The dominant role of low-frequency interaural time difference in sound localization

- tion[J] The Journal of the Acoustical Society of America, 1992, 91(3): 1648–1661.
- [19] Rui Y Q, Yu G Z, Xie B S, et al. Calculation of individualized near-field head-related transfer function database using boundary element method[C]. Audio Engineering Society 134th Convention, 2013.
- [20] Takeuchi T, Nelson P A, Kirkeby O, et al. Influence of individual head related transfer function on the performance of virtual acoustic imaging systems[C]. Audio Engineering Society 14th Convention, 1998.
- [21] Nelson P A, Bustamante O F, Engler E, et al. Experiments on a system for synthesis of virtual acoustic sources[J]. The Journal of the Audio Engineering Society, 1996, 44(11): 990–1007.
- [22] Liu L L, Xie B S. Analysis and experiment on the limitations of static and dynamic transaural reproduction with two frontal loudspeakers[J]. Archives of Acoustics, 2021, 46(2): 213–228.
- [23] Bai M R, Lee C C. Objective and subjective analysis of effects of listening angle on crosstalk cancellation in spatial sound reproduction[J]. The Journal of the Acoustical Society of America, 2006, 120(4): 1976–1989.