◇ 研究报告 ◇

# 双扬声器近场声源重放实验研究\*

魏 彤<sup>1,2</sup> 吴彦琴<sup>1,2</sup> 桑晋秋<sup>1,2†</sup> 郑成诗<sup>1,2</sup> 李晓东<sup>1,2</sup>

(1 中国科学院声学研究所 北京 100190)

(2 中国科学院大学 北京 100049)

**摘要:** 扬声器空间声重放系统的研究一般集中在重放声源的方向感方面,在距离感方面研究较少。该文采用 不同距离的近场头相关传递函数调整预滤波器系数,实现近场不同距离下声源的重放,研究多种典型双扬声 器配置下,等距离近场声源方向辨别和同方向近场声源距离辨别。实验结果表明:近场声源重放存在前后混淆 现象,当虚拟声源与扬声器同前方或同后方时,方向辨别较准确;在方向角判断准确的区域,能够较准确地分 辨声源的不同距离;声源越偏离中垂面,距离判定结果越准确。该研究为多扬声器配置下虚拟声源的距离重放 性能提供了较全面的评估。

关键词: 近场声源重放;双扬声器系统;近场头相关传递函数;串声消除

中图法分类号: O429 文献标识码: A 文章编号: 1000-310X(2020)02-0178-11 DOI: 10.11684/j.issn.1000-310X.2020.02.003

# Experimental study on near-field sound source reproduction using dual loudspeaker

WEI Tong<sup>1,2</sup> WU Yanqin<sup>1,2</sup> SANG Jinqiu<sup>1,2</sup> ZHENG Chengshi<sup>1,2</sup> LI Xiaodong<sup>1,2</sup>

(1 Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China)

(2 University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: The researches of louderspeaker spatial sound reproduction system mainly concentrate on the direction perception of the reproduced sound source while neglect the distance perception of the sound source. The near-field head related transfer functions (HRTFs) at different distances are used to adjust the prefilter coefficients, so as to reproduce near-field sound sources at different distances. The direction discrimination at equal distances and the distance discrimination at the same direction of the near field source is studied under several typical dual-loudspeaker configurations. The experimental results show that back-front confusion exists in such near-field sound reproduction, and the direction judgement is relatively more accurate, when the virtual sound source and the loudspeakers locate in the same front or back half-plane. In the area where the directions are accurately judged, the subject can differentiate the distances of the sound sources accurately, and the more the sound source deviates from the median plane, the more accurately the distance perception. This study provides a more comprehensive evaluation of the distance reproduction performance of virtual sound sources with multi-speaker configuration.

**Keywords:** Near-field sound source reproduction; Dual-loudspeaker system; Near-field head related transfer function; Cross-talk cancellation

<sup>2019-05-08</sup> 收稿; 2019-09-11 定稿

<sup>\*</sup>国家自然科学基金项目 (11504404, 11604362, 61801468), 中国科学院声学研究所青年人才项目 (QNYC201720)

作者简介:魏彤(1992-),男,甘肃会宁人,硕士研究生,研究方向:声信号处理。

<sup>†</sup>通信作者 E-mail: sangjinqiu@mail.ioa.ac.cn

# 0 引言

多通道扬声器声重放技术通过适当设计扬声 器布置和信号馈给而在一定的区域内产生期望的 听觉感知,以满足听音者临场感听音需求。该技术 主要包括基于人耳感知特性的虚拟声场重建和基 于声场合成特性的物理声场重建,目前已广泛应用 于电影院、音乐厅和会议室等诸多场合,其中有代 表性的矩阵环绕声技术包括5.1、7.1、9.1、22.2通路 系统等。分析表明,目前多通道扬声器声重放技术 主要侧重于声音的方向感、包围感、音色等,对于声 音的距离感仍缺乏研究。

现有少量研究通过数量庞大的扬声器实现对 近场声源的物理声场重建。Spors等<sup>[1]</sup>采用频谱相 除法(Spectral division method, SDM)利用直线形 扬声器阵列重放近场声源,但是该方法无法合成全 方位角的近场声源。随后,基于基尔霍夫-亥姆霍 兹方程的波场合成(Wave field synthesis, WFS)技 术和基于球谐函数的近场补偿的HOA(Near-field compensated higher order ambisonics, NFC-HOA) 技术采用圆形扬声器阵列来合成全方位角的近场 声源。然而,此类方法要求扬声器数量庞大,成本高, 复杂度高,且要求放音房间的混响时间短,在实际放 音环境下难以达到理想的效果。

如何利用较少数量的扬声器实现近场声源重 放,可基于人耳感知特性来实现。心理声学研究 结果表明,对于较远距离的声源(距离受试者超过 1 m), 受试者对距离感知主要依靠单耳因素, 包 括幅度、直达混响能量比及一些角度因素和谱因 素<sup>[2]</sup>;对于较近距离的声源(距离受试者小于等于 1 m), 受试者的躯干、肩、头等身体结构对波阵面产 生严重的影响,在偏离中垂面的方向,会引起低频 段很高的双耳声级差(Interaural level differences. ILDs)<sup>[3]</sup>,并且为受试者提供了近场声源距离感知 的因素<sup>[4]</sup>。而头相关传递函数(Head related transfer function, HRTF)能够较完整反映上述的单耳因 素和双耳因素。基于 HRTF 的双耳声重放技术可通 过HRTF 数据库合成双耳信号, 在双耳处实现声压 信号重建<sup>[5-7]</sup>。早期HRTF的测量主要基于远场条 件下,主要包括1994 年测量的MIT 库和2000 年 测量的CIPIC 库<sup>[8-9]</sup>。针对HRTF测量复杂和耗 时的问题,董秋洁等<sup>[10]</sup>提出了一种利用双扬声器 的HRTF动态测量方法,但对HRTF的近场测量数 据库较少; 1998年, Duda等<sup>[11]</sup>使用Bose Acoustimass 扬声器声源近似点源完成了近场HRTF的测 量,然而普通扬声器在近场条件下无法满足点源 条件; 2004年, Nishino等<sup>[12]</sup>使用十二面体扬声器 声源近似点源完成了近场HRTF的测量,但是十二 面体扬声器近似点源时,在一些频率范围内存在较 低的信噪比的问题; 2009年,中国科学院声学研究 所和北京大学合作采用KRMAR人工头测量了近 场HRTF数据库<sup>[13]</sup>,该数据库利用特殊的电火花 作为激励声源,具有较好的全指向性和高信噪比; 2018年,华南理工大学发布了不同距离的个性化近 场HRTF数据库<sup>[14]</sup>。

基于HRTF的双耳声重放技术具有实现简单、 计算复杂度低等优点。但该双通道信号通过双扬声 器重放时,双耳处声场存在串音,从而模糊虚拟声像 的方位信息,干扰聆听者对虚拟声像的判断,降低重 放效果<sup>[15]</sup>。针对这一问题, 1961年, Bauerd<sup>[16]</sup>提 出了串声消除系统(Cross-talk cancellation system, CCS); 1963年, Schroeder 等<sup>[17]</sup>将串声消除系统实 现并应用到研究之中。双扬声器系统的串声路径可 看作一个两输入两输出的系统,用2×2的空间响应 传递矩阵表示。相应的处理串音干扰的串音消除 器(Cross-talk canceller, CTC)是直接测量或者建 模得到的空间响应传递矩阵的逆<sup>[17-18]</sup>,对输入扬 声器之前的信号进行预滤波处理,从而消除系统的 串音干扰。对基于串声消除技术的虚拟立体声声场 重放, Ward 等<sup>[19-21]</sup> 使用参数加权的方法扩大了 声场甜点区范围,提高其鲁棒性;Kirkeb等<sup>[22]</sup>引入 了最小二乘的方法求解CTC,同时提出了声偶极 子 (Dipole) 布放方式; Bai 等<sup>[23-25]</sup> 针对 CCS 在扬 声器阵列和手机等实际重放系统中的应用,进行了 大量实验性的工作。但已有的双扬声器串声消除系 统着重研究远场声源的方向感、环绕感,尚未涉及 近场声源的重放及距离感研究。

本文以串声消除原理为理论基础,展开双扬声 器配置下近场声源不同方向、不同距离条件下虚拟 声源重放实验研究。研究了多种典型双扬声器配置 下的近场虚拟声像的方向感知和距离感知效果,包 括 Dipole 配置、立体声配置、环绕声配置以及扬声 器分别位于双耳两侧以考察前后混淆现象。实验 分两部分依次展开:实验一用于测试近场距离下的 虚拟声像方向辨别准确性,分析方向辨别准确区域 和方向辨别模糊区域;实验二排除方向辨别模糊区 域,在方向辨别准确的区域内,测试分析近场虚拟声像的距离辨别效果。实验结果表明,在方向角判断 准确的区域,受试者能够较准确地分辨声源的不同 距离,同时声源在侧方时受试者的距离判定结果更 准确。

# 1 虚拟声源的重放模型

# 1.1 模型建立

基于串声消除的双扬声器重放系统,其实质是 改变扬声器的信号输入,使得扬声器重放时的双耳 声压等于单声源的情况,也称这种方法为听觉传输 方法<sup>[26]</sup>。因此,双扬声器系统近场声源的重放是通 过调节两路扬声器信号的馈给,使得扬声器重放的 双耳声压等于距离受试者1m范围以内的单声源 的情况。图1给出了双扬声器系统的近场声源重放 的示意图,声源信号S经过两个单输入单输出的系 统,将幅度和相位调节后的信号输入到左右两扬声 器播放,最后信号到达受试者双耳产生与声源信号 S在近场情况下相同的效果。假设声源信号S经过 预滤波器 *H*<sub>pre</sub>处理后形成两扬声器的输入信号*X*, 可表示为

$$\boldsymbol{X}(e^{j\omega}) = \boldsymbol{H}_{pre}(e^{j\omega})S(e^{j\omega}), \qquad (1)$$

其中,  $\boldsymbol{H}_{\text{pre}}(e^{j\omega}) = \begin{bmatrix} H_{\text{L}}(e^{j\omega}) \\ H_{\text{R}}(e^{j\omega}) \end{bmatrix}$ ,  $H_{\text{L}}(e^{j\omega})$ 和

 $H_{\rm R}(e^{j\omega})$ 分别表示左右预滤波器。为了公式简洁, 后面的推导省略式中 $(e^{j\omega})$ 。





Fig. 1 Diagram of near field source reproduction with dual speaker system

信号 X 经过扬声器系统(系统包括声卡、音箱) 形成重放信号 Y,假设扬声器系统的频率响应为 H<sub>0</sub>,则此过程可以表示为

$$\boldsymbol{Y} = \boldsymbol{H}_0 \boldsymbol{X}, \tag{2}$$

其中,  $H_0 = \begin{bmatrix} H_{0L} & 0 \\ 0 & H_{0R} \end{bmatrix}$ ,  $H_{0L}$  和  $H_{0R}$  分别表示左 右扬声器系统的频率响应。

扬声器重放信号Y经空间传播后到达受试者 双耳,形成双耳信号E。则数学形式可表示为

$$\boldsymbol{E} = \boldsymbol{H}\boldsymbol{Y},\tag{3}$$

其中,  $\boldsymbol{H} = \begin{bmatrix} H_{\text{LL}} & H_{\text{RL}} \\ H_{\text{LR}} & H_{\text{RR}} \end{bmatrix}$ , 表示扬声器到受试者双

耳的系统传递矩阵, H<sub>LL</sub>、H<sub>LR</sub>、H<sub>RR</sub>和 H<sub>RL</sub>分别表示从左右扬声器位置到受试者左右耳的传递函数。

另一方面,假设空间中某一单声源位于图中虚 拟声源V的位置,产生信号S经空间传播后到达受 试者双耳,形成双耳信号**E**′。则数学形式可表示为

$$\boldsymbol{E}' = \boldsymbol{H}_{\mathrm{V}}S,\tag{4}$$

其中, $H_{V} = \begin{bmatrix} H_{VL} \\ H_{VR} \end{bmatrix}$ ,表示虚拟声源V到受试者双

耳的系统传递矩阵, H<sub>VL</sub>、H<sub>VR</sub>分别表示虚拟声源 位置到受试者左右耳的传递函数。

要使得扬声器系统重放的声像与单声源产生 相同的听觉效果,则需满足条件 *E'* = *E*。根据上 述条件,将式(1)、式(2)代入式(3),并结合式(4)可 以得到

$$\boldsymbol{H}\boldsymbol{H}_{0}\boldsymbol{H}_{\mathrm{pre}}\boldsymbol{S} = \boldsymbol{H}_{\mathrm{V}}\boldsymbol{S}.$$
 (5)

在实际应用中,扬声器系统的频率响应通常会 做均衡处理,从而实现两扬声器系统的理想匹配。 因此,在本文的研究中扬声器系统的频率响应 *H*<sub>0</sub> 可以忽略,则式(5)可简化为

$$\boldsymbol{H}\boldsymbol{H}_{\mathrm{pre}}\boldsymbol{S} = \boldsymbol{H}_{\mathrm{V}}\boldsymbol{S}.$$
 (6)

此时,预滤波器 Hpre 可以表示为

$$\boldsymbol{H}_{\rm pre} = \boldsymbol{H}^{-1} \boldsymbol{H}_{\rm V},\tag{7}$$

其中,**H**<sup>-1</sup>表示扬声器到受试者双耳的系统传递矩阵的逆矩阵。

实际中布放的扬声器离受试者的距离一般在 1.5 m以上(满足远场条件),因此需要选择远场的传 递函数 $H_{LL}$ 、 $H_{LR}$ 、 $H_{RR}$ 和 $H_{RL}$ 。根据重放虚拟声源 的位置信息,选择近场的传递函数 $H_{VL}$ 和 $H_{VR}$ 。根 据式(7)可以得到预滤波器 $H_{pre}$ 。

#### 1.2 模型求解

预滤波器 H<sub>pre</sub>的求解,核心问题是对系统传递 矩阵 H 进行求逆,一般采用伪逆法。即使使用伪逆 法,在某些频率范围也会出现病态问题,通常在求逆 过程中加入归一化代价函数(Tikhonov 正则化),使 得逆矩阵逼近真实的系统传递矩阵的逆。那么,系 统传递矩阵 H 的逆矩阵可以表示为<sup>[22,27]</sup>

$$\boldsymbol{H}^{-1} = \left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \beta\boldsymbol{I}\right]^{-1}\boldsymbol{H}^{\mathrm{H}}, \qquad (8)$$

其中,  $H^{H}$ 表示系统传递矩阵 H 的转置矩阵,  $\beta$ 表 示正则化代价因子, 用于限制滤波器增益和补偿奇 异值以保证该矩阵可逆, 该参数可为常数或与频率 相关。如果 $\beta$ 取值太小, 式(8)的结果 $H^{-1}$ 在某些 频点上出现尖锐的峰值; 反之, 式(8)的结果相比于 理想结果存在较大误差。在实际的求逆过程中, 一 个精确的 $\beta$ 值不是至关重要<sup>[22]</sup>, 但是 $\beta$ 要保证足够 的通道分离度, 即串声与直达声之间的幅度比足够 小<sup>[27]</sup>。

本文取 $\beta = 0.005$ 作为求解逆矩阵过程中的正则化因子,客观结果表明通道分离度足够,虽然引入  $\beta$ 值会带来信号音色的轻微改变,但能满足受试者 测听定位实验的要求。接着将式(9)代入式(7)可求 得预滤波器的形式为

$$\boldsymbol{H}_{\text{pre}} = \left[\boldsymbol{H}^{\text{H}}\boldsymbol{H} + \beta\boldsymbol{I}\right]^{-1}\boldsymbol{H}^{\text{H}}\boldsymbol{H}_{\text{V}}.$$
 (9)

将式(9)得到的结果 **H**<sub>pre</sub>代入式(1)可求得声源信 号 S 经过预滤波器之后的信号 **X**,即为扬声器的输 入信号。

# 2 测听实验

#### 2.1 实验配置

实验在中国科学院声学研究所的音频声学测 听室中完成,室内装饰吸声材料和吸声板,尺寸为 4.22 m×4.05 m×3.28 m,混响时间约为0.16 s,本底 噪声平均值为14.8 dB(A)。如图2所示,实验使用 的扬声器为两只Genelec 8030B有源音箱,声卡选 用 Antelope Orion 32,其中扬声器离受试者的距离 为1.5 m,且声卡与扬声器之间使用专业Canare音 频线连接。实验中模拟声源的信号为单通道 click 信 号 (图3为信号片段的短时傅里叶功率谱),信号频 率主要在2 kHz,采样率为44.1 kHz。

在双耳处输入信号的计算中,选择 MIT 数据 库<sup>[8]</sup>中的 HRTF 作为左右两扬声器到受试者双耳 的传递函数,来代表传递函数 H<sub>LL</sub>、H<sub>LR</sub>、H<sub>RR</sub>和 H<sub>RL</sub>;选择 IOA-PKU数据库<sup>[13]</sup>中的 HRTF 来代表 虚拟声源到受试者双耳的传递函数 H<sub>VL</sub>、H<sub>VR</sub>,最 后根据式(9)求得预滤波器。将 click 信号通过预 滤波器的处理,可以得到左右扬声器的输入信号 X。两数据库在测量 HRTF 时都是在全消声室采用 kemar 人工头进行测量,头部的衍射和耳廓耳道的 声滤波特性近似一致,可以避免近远场数据库由于 测量条件的较大差异带来的负面影响。但由于两数 据库测量环境、测量条件和测量方法存在一定差异, 而且没有采用个性化的 HRTF,近远场数据库代入 公式(9) 后仍然会带来一定的感知误差。



图2 音频声学测听实验室





# 2.2 实验设计

测听实验使用上文描述的测听环境和实验器材。听力正常的6位受试者(3男3女)参加了测听实验。根据双扬声器的夹角大小展开了不同的扬声器 配置模式下的测听实验:

(1)扬声器夹角为10°(声偶极子);

(2)扬声器夹角为60°(标准立体声);

(3)扬声器夹角为180°(扬声器与中垂面的夹 角为直角);

2020年3月

(4) 扬声器夹角为220°(标准5.1 声道中的环绕 声道)。

实验分两部分依次展开,分别进行对虚拟声源 方位角和距离的判定。

实验一安排受试者进行虚拟声源方位角的判定,以确定虚拟声源方向判定的准确区域和模糊区域。虚拟声源方位角的判定在虚源距离等于20 cm、75 cm两种情形下完成。图4给出了在4种不同的扬声器配置模式下需要受试者辨别的虚拟声源的方向(用虚线表示)。当扬声器夹角为10°时,如图4(a)所示,要求每位受试者辨别水平面全方位采样间隔为30°的虚拟声源的方向。在此过程中,实验员随机重放图4(a)中虚线所示的不同方向的虚拟声源,并要求受试者将感受到声源的方向与图4(a)中虚线所示的12个方向对应并反馈给实验员,实验员随即记录反馈结果。

根据初步的测听实验结果,图4(b)~ 图4(d) 三 种配置下受试者对在左半平面和右半平面的虚源 感知定位具有近似对称性。因此为了减少受试者的 疲劳感对实验结果的影响,图4(b)~ 图4(d) 三种配 置下要求每位受试者完成右半平面采样间隔为15° 的虚拟声源的方向角判定。在此过程中,实验员随 机重放上述不同方位的虚拟声源,并要求受试者将 感受到声源的方向与图4(b)~ 图4(d)中虚线所示 的13个方向对应并反馈给实验员,实验员随即记录 反馈结果。要求每位受试者连续完成虚拟声源距 离等于20 cm、75 cm两种情形下的测听,在此过程 中,受试者每完成一种距离情形下的测听之后,安排 5 min时间的休息。当受试者完成实验一虚拟声源 方位角的判定之后,安排15 min时间的休息,以便 排除测听疲劳对后续实验过程的影响。





Fig. 4 Schematic diagram of loudspeaker configuration and virtual source azimuth to be measured for Experiment One

实验二安排受试者进行声源距离的判定。在实验中,受试者可以用声源在肩部、肘部或手部去衡量感知的声源距离,实验员随后记录判定结果,根据每位受试者的肩部、肘部和手部的实际距离将判定结果等效为测量距离。这种方式对距离的判定更精确,因为没有限定受试者必须把感知距离判定成2 cm、50 cm 和75 cm 几个离散距离选项。因此,在实验之前,需要完成对每位受试者的肩、肘、手3个位置的距离的测量。表1给出参加实验的受试者肩、肘、手3个位置的测量距离。

表1 受试者肩部、肘部、手部距离的测量结果 Table 1 The shouder length, elbow length and hand length

受试者序号	肩部距离/cm	肘部距离/cm	手部距离/cm
受试者 A(女)	19	45	74
受试者 B(女)	17	43	72
受试者 C(男)	25	48	80
受试者 D(女)	24	49	79
受试者 E(男)	23	45	82
受试者 F(男)	23	47	72

根据实验一受试者对虚拟声源方位角的判定 结果,当扬声器夹角为180°(图4(c))时,受试者出 现严重的前后混淆现象(如图6(e)、图6(f)所示)。因 此,在实验二测听辨别距离时,只测听另外3种配 置下受试者对虚拟声源距离的感知,即扬声器之间 夹角为10°、60°和220°三种情况(图5)。图5给出 了在3种不同的扬声器配置模式下方位角判断准确 的区域内需要受试者判断的虚拟声源的距离(用× 表示)。当扬声器夹角为10°时,如图5(a)所示,在 虚拟声源方位角为0°、30°、60°、300°和330°五种情 形下,要求受试者感知实际距离为20 cm、50 cm和 75 cm的虚拟声源的距离。当扬声器夹角为60°时, 如图 5(b) 所示,在虚拟声源方位角为0° 和 30° 的情 形下,要求受试者感知实际距离为20 cm、50 cm 和75 cm的虚拟声源的距离。当扬声器夹角为220° 时,如图5(c)所示,在虚拟声源方位角为120°、150° 和180°三种情形下,要求受试者感知实际距离为 20 cm、50 cm 和 75 cm 的虚拟声源的距离。图 5(b) 和图5(c)两种扬声器配置模式在预实验中的声像 感知效果呈左右近似对称性,且为减少受试者疲劳, 实验仅测试了虚源在右侧的区域。



图5 实验二扬声器配置模式及虚源待测距离示意图

Fig. 5 Schematic diagram of loudspeaker configuration and virtual source distance to be measured for Experiment Two

# 3 实验结果分析

# 3.1 方位角判定结果分析

在扬声器夹角不同的情况下,图6根据6位受 试者的测听结果分别绘制了虚源距离为20 cm 和 75 cm时的角度辨别散点图。如图6所示,横轴表示 目标方位角,纵轴表示判断方位角,虚线表示方位角 判断的理想结果,圆点的大小代表所有受试者角度 判断结果的重合次数。圆点越大判定结果重合的次 数越多,也表明受试者对该方位的虚拟声源的角度 感知结果越一致。

图 6(a) 和图 6(b) 分别给出了受试者在扬声器 夹角为10°时,对距离为20 cm 和75 cm 的虚拟声源 方位角的测听结果。从两图中可知,大部分受试者 可以准确判定前半水平面的虚拟声源,但对于正前 方和正后方的虚拟声源,受试者存在严重的前后混 淆现象。对于右侧区域的虚拟声源,受试者基本将 其判定为30°或者60°方向的虚拟声源,但似乎结果 存在较大的离散性(例如最为明显的120°方向,散 点分散度很高)。然而,对于左侧区域的虚拟声源, 受试者完全将其判定为300°或者330°方向的虚拟 声源,结果集聚性很强。这说明偶极子布放方式重 放的声像左侧区域比右侧区域明显。 图 6(c) 和图 6(d) 分别给出了受试者在扬声器 夹角为 60° 时,对距离为 20 cm 和 75 cm 的虚拟声源 方位角的测听结果。从两图中可知,当重放虚拟声 源位于前半水平面时,受试者基本都能准确判定其 位置。当重放虚拟声源位于后半水平面时,受试者 将实际的后方虚拟声源刘定为前方虚拟声源,出现 后前混淆问题。当重放虚拟声源位于扬声器夹角区 域之外时,受试者很难区别这个区间的虚拟声源的 具体方向。因此,本文称这个区间为方向辨别模糊 区间。比较两图可知,当虚拟声源距离增大时,受试



Fig. 6 Results of the azimuth localization in Experiment One

者的方向辨别模糊区间变小。

图6(e)和图6(f)分别给出了受试者在扬声器夹 角为180°时,对距离为20 cm和75 cm的虚拟声源 方位角的测听结果。从两图中可知,当重放虚拟声 源位于前半水平面时,受试者前后混淆问题严重。 当重放虚拟声源位于后半水平面时,混淆问题减弱。 其中重放虚拟声源位于90°时,受试者都能准确判 定其方位角。比较两图可以看出,当虚拟声源距离 增大时,受试者的测听结果的准确度提高。

图 6(g) 和图 6(h) 分别给出了受试者在扬声器 夹角为220°时, 对距离为20 cm 和75 cm 的虚拟声 源方位角的测听结果。从两图中可知, 当重放虚拟 声源位于后半水平面时, 受试者基本都能准确判定 其位置。当重放虚拟声源位于前半水平面时, 受试 者基本完全将前方虚拟声源判定为关于前平面对 称的后方虚拟声源, 出现前后混淆问题。比较两图, 当虚拟声源距离增大时, 受试者的测听结果准确度 提高。

对于前后混淆现象, 文献 [6] 中指出, 耳廓对声 波的反射和散射所引起的谱因素对前后方向的声 源定位非常重要。但耳廓引起的谱因素极具个性 化特征,且仅对高于5kHz~6kHz的高频声波有 效,而且由于头部偏移耳廓难以带来稳定可靠的 定位因素。而对于1kHz~2kHz以下的低频情况, HRTF是近似前后对称的,难以分清前后。由于上 述原因非个性化HRTF极易产生前后混淆现象。

总之,当虚源与扬声器同处于前半水平面或同 处于后半水平面时,受试者可以准确判定虚拟声源 的方向;当重放虚拟声源位于扬声器夹角之外时,受 试者无法准确判定虚拟声源的方向;当虚拟声源距 离增大时,受试者的测听结果准确度提高。

### 3.2 距离判定结果分析

在扬声器夹角为10°、60°和220°的情况下, 图7~图9根据受试者对于不同方向虚拟声源距离 的感知结果绘制了散点图。其中,横轴表示目标距 离,纵轴表示判定距离,虚线表示距离判定的理想结 果,圆点代表每位受试者对虚拟声源距离的判定结 果。图7~图9显示的主观距离判定是受试者感知 的实际结果,而不是把距离判定成离散的某几个目



图7 扬声器夹角10°,受试者虚源距离测听结果

Fig. 7 The angle between the two speakers is  $10^{\circ}$ , and the results of the distance localization

在用声学



图8 扬声器夹角60°,受试者虚源距离测听结果

Fig. 8 The angle between the two speakers is  $60^{\circ}$ , and the results of the distance localization







标距离,一定程度上可以更准确地衡量受试者的实际距离感知。由于不同受试者肩部、肘部、手部的测量长度存在相同的情况,因此对于感知距离结果重合的情况采用较大的圆点表示。

当扬声器夹角为10°时,距离感知散点图如图7 所示。其中,图7(a)表示虚源位于0°方向时不同受 试者虚源距离的判断结果,从图中可看出受试者对 0°方向虚拟声源的距离感较差。当虚拟声源的方向 偏离0°时,受试者对不同目标距离的虚拟声源定位 准确度明显提高,其中最明显的为侧方声源,即60° 方向和300°方向。将右侧与左侧的测听结果对比, 可以看出受试者对左侧虚拟声源的感知距离较近。

当扬声器夹角为60°时,距离感知散点图如图8 所示。其中,图8(a)表示虚源0°方向时不同受试者 虚源距离的判断结果,从图中可看出受试者对0° 方向虚拟声源的距离感较差,所有受试者都存在多 次判断错误的情况,尤其当虚拟声源的目标距离为 20 cm时,受试者完全无法正确判定。图8(b)表示 虚源30°方向不同受试者虚源距离的判断结果,从 图中可看出受试者对30°方向虚拟声源的距离判定 结果分布更靠近图中理想虚线两侧,大部分受试者 可以正确感知目标距离为20 cm处的虚拟声源,所 有受试者可以正确感知目标距离为75 cm的虚拟声 源。总体而言,受试者对稍远距离声源和侧方声源 的距离感知更准确。

当扬声器夹角为220°时,距离感知散点图如 图9所示。其中,图9(a)表示虚源120°方向时不同 受试者虚源距离的判断结果,从图中可看出受试者 对120°方向的虚拟声源距离感较好,判定结果分布 靠近图中理想虚线两侧;图9(b)表示虚源150°方向 时不同受试者虚源距离的的判断结果,从图中可看 出大部分受试者可以正确判断不同距离的虚源,相 比120°方向虚拟声源的判定结果,其结果更偏离于 理想虚线两侧;图9(c)表示虚源180°方向时不同受 试者虚源距离的判断结果,相比于120°和150°方向 虚拟声源的判定结果,图9(c)中受试者的判定结果 最偏离于理想虚线两侧,受试者存在多次判断错误 的情况。因此,受试者对侧方虚拟声源的距离感优 于180°方向虚拟声源的距离感,其中对120°方向 的虚拟声源的距离感最好,基本可准确区分不同距 离的虚源。同样说明受试者对侧方声源的距离感更 准确。

## 4 结论

本文基于不同距离的HRTF,通过串声消除的 方法实现在多种典型的双扬声器配置下近场虚拟 声源的重放,并对其完成了实验测听研究。测听结 果表明该双耳合成技术存在前后混淆现象,一对前 方左右扬声器布置难以产生后半水平面范围的虚 拟源,一对后方左右扬声器不能产生前半水平面的 虚拟源。但在方位角判断准确的方向上,可对虚拟 源感知距离进行控制,特别是当方向角越偏离中垂 面时,对感知距离控制的准确度越高。该测听结果 也验证了所采用的近场HRTF数据库的有效性。如 何在双扬声器重放近场声源时增强距离控制准确 度也值得进一步研究。



- Spors S, Ahrens J. Reproduction of focused sources by the spectral division method[C]. 2010 4th International Symposium on Communications, Control and Signal Processing (ISCCSP). IEEE, 2010: 1–5.
- [2] Zahorik P, Douglas S B, Adelbert W B. Auditory distance perception in humans: a summary of past and present research[J]. Acta Acustica united with Acustica, 2005, 91(3): 409–420.
- [3] Brungart D S, Rabinowitz W M. Auditory localization of nearby sources. Head-related transfer functions[J]. Journal of the Acoustical Society of America, 1999, 106(3Pt1): 1465–1479.
- [4] Brungart D S. Auditory localization of nearby sources.
   III. Stimulus effects[J]. Journal of the Acoustical Society of America, 1999, 106(6): 3589–3602.
- [5] Hammershøi D, Møller H. Binaural technique—Basic methods for recording, synthesis, and reproduction[M]// Communication Acoustics. Berlin, Heidelberg: Springer, 2005: 223–254.
- [6] 谢菠荪. 头相关传输函数与虚拟听觉 [M]. 北京: 国防工业出版社, 2008.

- [7] Bauer B B. Stereophonic earphones and binaural loudspeakers[J]. Journal of the Audio Engineering Society, 1961, 9(2): 148–151.
- [8] Gardner W G, Martin K D. HRTF measurements of a KEMAR[J]. Journal of the Acoustical Society of America, 1995, 97(6): 3907–3908.
- [9] Algazi V R, Duda R O, Thompson D M, et al. The cipic hrtf database[C]. Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575). IEEE, 2001: 99–102.
- [10] 董秋洁,叶青华,李晓东.利用双扬声器的头相关传递函数动态测量 [J]. 电声技术, 2011, 35(2): 54-56.
  Dong Qiujie, Ye Qinghua, Li Xiaodong. Dynamic head-related transfer function measurement using dual-loudspeaker array[J]. Audio Engineering, 2011, 35(2): 54-56.
- [11] Duda R O, Martens W L. Range dependence of the response of a spherical head model[J]. Journal of the Acoustical Society of America, 1998, 104(5): 3048–3058.
- [12] Nishino T, Hosoe S, Takeda K, et al. Measurement of the head related transfer function using the spark noise[C].
  18th International Congresses on Acoustics (ICA2004) Tu5. D, 2004, 1: 1437–1438.
- [13] Qu T, Xiao Z, Gong M, et al. Distance- dependent head-related transfer functions measured with high spatial resolution using a spark gap[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2009, 17(6): 1124–1132.
- [14] Yu G, Wu R, Liu Y, et al. Near-field head-related transferfunction measurement and database of human subjects[J]. Journal of the Acoustical Society of America, 2018, 143(3): EL194–EL198.
- [15] Bauck J, Cooper D H. Generalized transaural stereo and applications[J]. Journal of the Audio Engineering Society, 1996, 44(9): 683–705.
- [16] Bauer B B. Stereophonic earphones and binaural loudspeakers[J]. Journal of the Audio Engineering Society, 1961, 9(2): 148–151.
- [17] Schroeder M R, Atal B S. Computer simulation of sound transmission in rooms[C]. Proceedings of the IEEE, 1963, 51(3): 536–537.
- [18] Damaske P. Headrelated twochannel stereophony with loudspeaker reproduction[J]. Journal of the Acoustical Society of America, 1971, 50(4B): 1109–1115.
- [19] Ward D B, Elko G W. Optimum loudspeaker spacing for robust crosstalk cancellation[C]. Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181). IEEE, 1998, 6: 3541–3544.
- [20] 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.
- [21] Ward D B. Joint least squares optimization for robust acoustic crosstalk cancellation[J]. IEEE Transactions on Speech and Audio Processing, 2000, 8(2): 211–215.

- [22] Kirkeby O, Nelson P A, Hamada H, et al. Fast deconvolution of multichannel systems using regularization[J]. IEEE Transactions on Speech and Audio Processing, 1998, 6(2): 189–194.
- [23] Bai M R, Tung C W, Lee C C. Optimal design of loudspeaker arrays for robust cross-talk cancellation using the Taguchi method and the genetic algorithm[J]. Journal of the Acoustical Society of America, 2005, 117(5): 2802–2813.
- [24] Bai M R, Lee C C. Development and implementation of cross-talk cancellation system in spatial audio reproduction based on subband filtering[J]. Journal of Sound and Vibration, 2006, 290(3–5): 1269–1289.
- [25] 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.
- [26] Bauck J, Cooper D H. Generalized transaural stereo and applications[J]. Journal of the Audio Engineering Society, 1996, 44(9): 683–705.
- [27] 许春冬, 李军锋, 裘嫄, 等. 两扬声器配置下的串声消除系统参数优化设置 [J]. 计算机应用, 2014, 34(5): 1503–1506, 1525. Xu Chundong, Li Junfeng, Qiu Yuan, et al. Parameters design and optimization of crosstalk cancellation system for two loudspeaker configuration[J]. Journal of Computer Applications, 2014, 34(5): 1503–1506, 1525.

◇ 声学新闻和动态 ◇

# 新型声学超材料单通道麦克风可实现多声源实时定位与分离

在人工系统中,科研人员通常借助由多个传声器组成 的传声器阵列来解决声源定位和分离问题。具有高精度声 源定位和分离能力的传声器阵列往往需要较大的阵元数量 和物理尺寸,这种阵列系统不仅不便于安装和操控,处理多 通道信号的计算成本往往也很大,从而导致其应用受限。

受生物听音机制的启发,中国科学院噪声与振动重 点实验室的博士生孙雪聪与其导师杨军研究员、贾晗研究 员等提出了一种基于声学超材料的单通道多声源的定位 与分离系统,用一个带有超材料外壳的单通道麦克风实现 了三维空间中多个同时发声声源的实时定位与分离。相关 研究成果已发表于国际学术期刊Advanced Science (DOI: 10.1002/advs.201902271),并被选为当期封面文章 (inside back cover)。

研究人员将麦克风嵌入到精心设计的三维超材料结构 中,该结构以与来波方向相关的方式修改麦克风的频率响 应,从而对来自三维空间中不同方向的信号进行编码。研究 人员还提出了一种联合重建算法VSPCA-OMP,该算法具 有较低的复杂度,可以基于采集到的单通道数据实现多声源 的实时定位和分离。

为了证明该系统的定位与分离能力,研究人员针对多 个场景开展了听音测试,结果表明,当空间中同时发声的声 源数量不超过3个时,该系统的定位与分离的平均准确率维 持在90%以上。由于所提出的算法复杂度较低,每次重建过 程耗时均控制在1s以内,良好的实时性使该系统也可用于 识别和追踪声目标。

未来该系统有望应用于智能场景监测、机器听觉及语 音识别的前端处理等领域。

(中国科学院声学研究所 孙雪聪)