DOI: 10.1051/jnwpu/20183620281

# 基于多维生理参数的头相关传递函数个人化方法

黄婉秋, 曾向阳 , 王蕾

(西北工业大学 航海学院, 陕西 西安 710072)

**摘 要:** 头相关传递函数(HRTF)是实现听觉空间虚拟的关键函数,不仅与声源的相对方位、频率有关,还与听者的生理结构有密切的关系,这使得其个人化(personalization)成为相关应用中的瓶颈问题。为此,以实测 HRTF 和多维生理参数数据为基础,结合相关分析、主成分分析和数据库匹配算法,给出了一种个人化方法,并通过谱失真(SD)以及主观听觉实验对算法的有效性进行评价和验证。结果表明,基于最小距离准则的数据库匹配方法是实现 HRTF 个人化预测的一种有效方法;基于多维实测生理参数获得的个人化 HRTF 数据相比于基于人工头的 HRTF 数据,能提高定位准确率 4.6%,降低前后混淆率 2.2%。

**关 键 词:** HRTF;个人化 ;生理参数 ;相关性分析 ;主成分分析 ;数据库匹配 ;听觉定位

**中图分类号:** TN912 **文献标志码:** A **文章编号:** 1000-2758(2018)02-0281-06

头相关传递函数(head-related transfer function, HRTF)是实现听觉空间虚拟的关键函数,在室内声场建模、听觉虚拟现实、3D 立体声等领域都具有重要的应用价值,也是声学、多媒体等学科的热门话题。头相关传递函数描述了声波从声源到双耳的传递过程,时域为头相关脉冲响应(head-related impulse response, HRIR)。它体现了人的生理结构(如头、耳廓以及躯干等)对声波的综合滤波效应,包含了有关声源的位置信息[1],是具有明显个人特征的物理量。在虚拟听觉重放的应用中,采用倾听者本人的 HRTF 将会得到较好效果,而采用其他 HRTF 的效果就取决于该数据与倾听者 HRTF 的相似性。HRTF 的测量是一项非常耗时并且复杂的工作,因此,如何快速有效地获得与任意听者生理参数相对应的 HRTF 数据(HRTF 个人化)已成为制约空间虚拟听觉技术的关键问题[2]。

国外已公布了一些人工头和真人测试数据,但一方面是公布的数据信息不够完整,另一方面考虑人种之间的差异,这些研究数据多数不适合中国人的生理特征,故而需要通过实验测量或理论计算来获取真实有效的 HRTF 数据。理论计算着重于求解,目前常用的理论计算方法有头部刚球模型[3]、雪人模型计算方法[4]以及 HRTF 数值计算方法如边界元法[5]。实验测量是获得 HRTF 最重要且最准确的手段,常采用封闭耳道测量方法。由于理论计算方法的计算量大,实验测量需要一系列专用的设备,且测量结果不稳定易受影响,因此单一的理论计算方法或实验测量方法目前还不足以同时满足高计算速度、宽频带和个人化的要求。

近年来,研究者试图采用近似方法建立个人化 HRTF 的快速获取方法,提出了生理参数匹配法[6]、频率标度变换法[7]、生理参数线性回归法[8]和结构模型法[9]、结合感知特性改善 HRTF 建模精度[10]和基于主成分分析法的个人化 HRTF 的获取[11]等方法。由于参数匹配法考虑的生理参数有限,还没有完全确定与 HRTF 密切相关的生理参数。

本文以 53 名成年中国人的多维实测生理参数和 HRTF 数据库(53 个真人对象的 723 个方位)作为数据基础,首先利用主成分分析法和相关性分析法优选出用于匹配的生理参数,再利用数据库匹配实现头相关传递函数的个人化,最后利用谱失真和主观听觉实验检验该方法的有效性。

**收稿日期:** 2017-04-01 **基金项目:** 国家自然科学基金(11774291)资助

**作者简介:** 黄婉秋(1994—),女,西北工业大学硕士研究生,主要从事室内声学和声信号处理研究。

# 1 头相关传递函数概念及其个人化

HRTF 是头部、躯干、耳廓等生理结构对声波综合作用的结果,即 HRTF 与生理参数存在相关性,因此,如果能设法获得那些相关性最强的独立参数,便可能通过建立模型近似估计或预测出受试者的 HRTF。基于此,本文实现 HRTF 个人化的研究方案如图 1 所示。首先对实测 53 个真人数据库的 HRTF 时域测量值(即头相关脉冲响应,HRIR)进行预处理,然后提取 HRTF 的主成分,同时将 50 维未经处理的生理参数进行自相关分析,初选若干参数,将其与 HRTF 的主成分进行互相关分析,从而得到用于匹配的关键生理参数,最后用这些生理参数进行数据库匹配得到受试者的近似 HRTF。再利用干信号与其构造虚拟声信号,组织相应的受试者进行主观听觉实验,对算法的有效性加以验证。

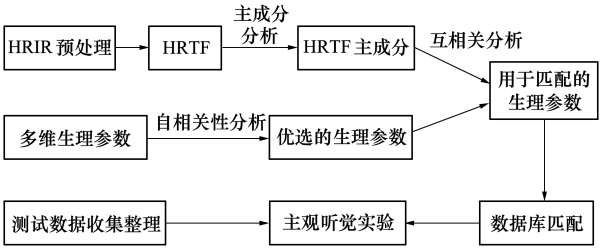


图 1 HRTF 个人化流程图

# 2 多维生理参数的优选

## 2.1 相关性分析

在选取生理参数的组合时,本文首先对实测的 53 人数据库中各个俯仰角、方位角的 HRIR 进行了预处理。由于虚拟听觉重放时主观音色会有改变,故需对 HRTF 进行均衡处理。利用信号源解卷积反变换后的结果与测试信号解卷积后的结果进行卷积得到了最终的 HRIR,为后续 HRTF 的主成分分析和相关性分析提供了数据基础。针对每个受试者测量的 50 维生理参数,首先利用自相关分析选取独立的参数,相关系数小于 0.5 认为不相关,对于相关系数取模后将小于 0.5 的相关系数归零得到自相关的灰度图(见图 2)。首先选取独立的生理参数,如图 2 中参数 2 容貌面长、参数 21 额最小宽等。再选取线性相关的一组中对 HRTF 影响更大且更易测量的参数[12]。例如参数 27 两耳屏间宽和参数 9 鼻尖枕突距相关,考虑到耳部参数对 HRTF 的影响更明显,故剔除参数 9 鼻尖枕突距,留下参数 27 两耳屏间宽。按此方法初步优选了 29 维生理参数,如图 4 所示。

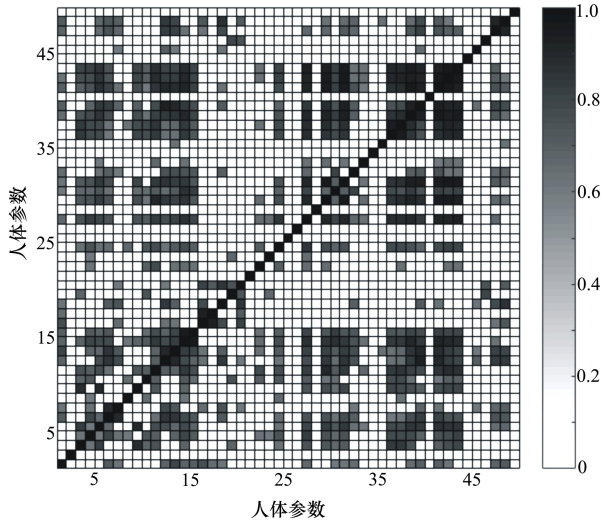


图 2 50 个生理参数的自相关分析

## 2.2 主成分分析

采用主成分分析对 HRTF 数据库各俯仰角的数据进行分析。步骤如下:

1) 根据已知的 $M$ 个空间方向的 $DTF$(每个方向 $N$ 个频率点),构造 $N \times M$ 矩阵 $\boldsymbol{H}_{\Delta N\times M}$

$$\boldsymbol{H}_{\Delta N\times M} = [H_{\Delta,0}(f) \quad H_{\Delta,1}(f) \quad \cdots \quad H_{\Delta,M-1}(f)] = \begin{bmatrix} H_\Delta(\theta_0,f_0) & H_\Delta(\theta_1,f_0) & \cdots & H_\Delta(\theta_{M-1},f_0)H_\Delta(\theta_0,f_1) \\ H_\Delta(\theta_1,f_{N-1}) & H_\Delta(\theta_{M-1},f_1) & \cdots & H_\Delta(\theta_{M-1},f_{N-1}) \end{bmatrix} \tag{1}$$

2) 求出该矩阵自协方差矩阵 $\boldsymbol{R}$,并按对应的本征值由大到小排序 $u_1,u_2,\cdots,u_N$,取前 $Q$ 个本征矢量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_Q$ 作为 $Q$ 个谱形状矢量,则 $Q$ 个谱形状矢量组成的矩阵为:

$$\boldsymbol{D}_{N\times Q} = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_Q] \tag{2}$$

3) 利用谱形状基矢量对 $\boldsymbol{H}_\Delta$ 进行分解,利用 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_Q$ 的正交性,可以求出与空间方向有关,但与频率无关的 $Q \times M$ 权重系数矩阵:

$$\boldsymbol{W}_{Q\times M} = \boldsymbol{D}_{Q\times N}\boldsymbol{H}_{\Delta N\times M} \tag{3}$$

4) 最后,将 $M$ 个空间方向的 HRTF 近似用 $Q$ 个谱形状基矢量表示:

$$\hat{\boldsymbol{H}}_{\Delta N\times M} = \boldsymbol{D}_{N\times Q}\boldsymbol{H}_{\Delta Q\times M} \tag{4}$$

进而有:

$$\hat{H}_i(f) = \hat{H}(\theta_i,f) = \sum_{q=1}^{Q} w_{qi}d_q(f) + H_{av}(f) = \sum_{q=0}^{Q} w_{qi}d_q(f) \tag{5}$$

因而 $H(\theta_i,f)$ 可分解为与方向无关的 $H_{av}(f)$ 以及 $Q$ 个谱形状基矢量的权重和,与谱形状基矢量对应的权重系数 $w_{qi}$ 的集合也称为主成分或空间特性函数,它决定了谱形状基矢量对不同方向 HRTF 的贡献。

前 $Q$ 个本征矢量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_Q$ 是正交的,但并不是完备的, $Q$ 取得越大(谱形状基矢量数目越多),近似程度越高,但数据越复杂。由于各谱形状基矢量对 $\boldsymbol{H}_\Delta$ 的贡献以 $\boldsymbol{u}_1$ 最大,$\boldsymbol{u}_2$ 次之,依次递减,因此,当 $Q$ 一定时,取 $N\times N$ 矩阵 $\boldsymbol{R}$ 的前 $Q$ 个本征矢量作为谱形状基矢量,就相当于取相对贡献最大的 $Q$ 个谱形状基矢量进行分解,所带来的平方误差最小[1]。

由各俯仰角的主成分分析的结果,统计分析得各个方位角的前 29 个主成分的贡献率都高于 80%,故根据上述方法分别取各俯仰角的前 29 个主成分,将其与 29 维生理参数分别进行互相关分析,对相关系数取模并将小于 0.5 的相关系数归零后得图 3,最后对各俯仰角互相关分析的结果进行统计和优选。综合考虑各俯仰角、各参数间的相关关系,最终优选了 12 维对 HRTF 影响显著的生理参数(见图 4)。

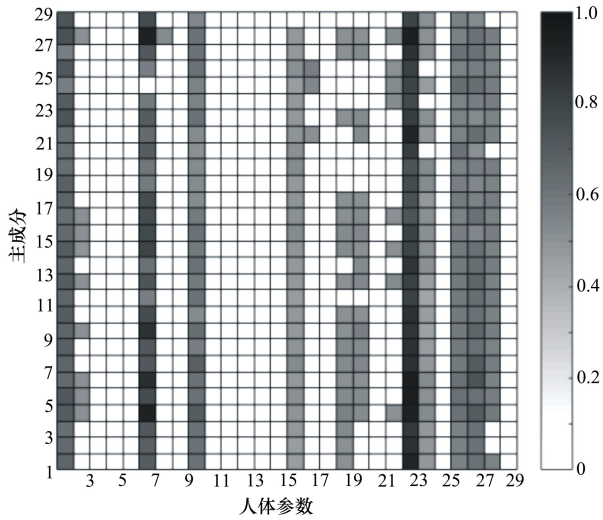


图 3　基于主成分分析的相关性分析图(俯仰角 $\phi=0^0$)

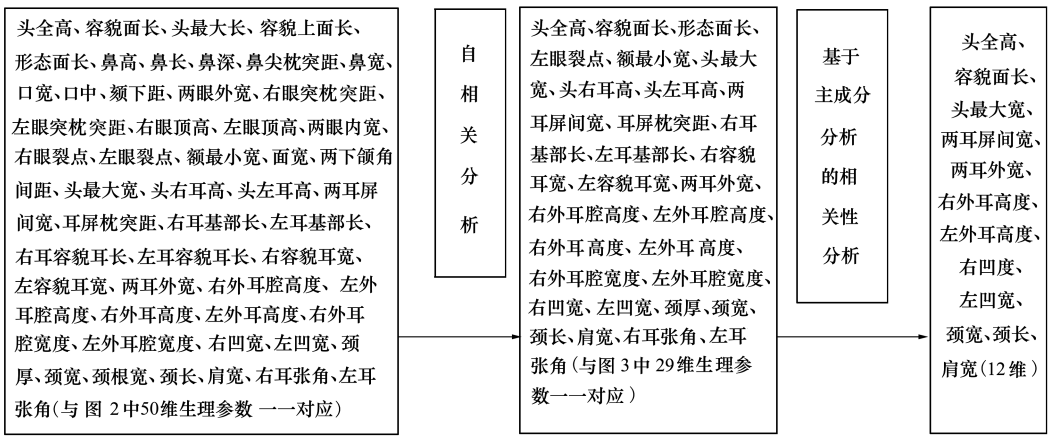


图 4　多维生理参数的优选

## 3　基于数据库匹配的 HRTF 个人化

本文的数据库匹配采用基于最小距离准则的数据库匹配算法,利用优选的 12 维生理参数进行数据库匹配。进行初步验证时,从数据库选取 5 名受试者,当选取其中一名受试者时,数据库中其他 52 名作为待匹配对象,根据匹配结果,确定与待预测听者最接近的已测 HRTF 的听者,认为二者的 HRTF 是近似相同。

对于某个空间方向,受试者个性化 HRTF 和用生理参数匹配法获取的近似 HRTF 之间的差异可采用谱失真 $SD$ 表示。$SD$ 在离散频率域的定义为

$$SD=\sqrt{\frac{1}{N}\sum_{f_k}20\lg\left(\frac{|H(f_k)|}{|H'(f_k)|}\right)^2}\qquad(6)$$

式中, $H'(f_k)$ 为新受试者个性化 HRTF 在 $N$ 个离散频率点 $f_k$ 的采样值,$H(f_k)$ 为通过匹配获得的近似 HRTF 在 $N$ 个频率点 $f_k$ 的采样值。$SD$ 值越小,说明匹配获得的 HRTF 与个性化 HRTF 偏离越小。

对比受试者和其匹配对象在各方位的结果,从曲线对比和谱失真两方面对匹配结果进行误差分析。图 5 为俯仰角为 0°,方位角为 45°匹配对象的

HRTF 与受试者的 HRTF。可以看出,2 条曲线整体趋势吻合,波峰波谷也大部分吻合,有些频率虽然幅值有差异,但趋势相同,显示匹配结果具有较好的准确性。

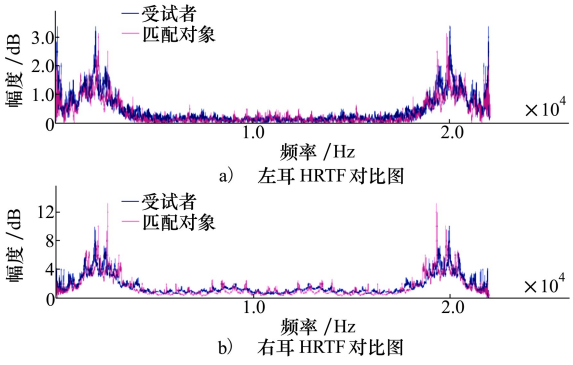


图 5　受试者与匹配对象 HRTF 对比图

对每个人在各俯仰角下的各象限分别取一方位角(即右前、右后、左前、左后方向各取一固定方位角),计算其谱失真,并对 5 人的谱失真取均值,结果如图 6 所示。可以看出谱失真($SD$)均小于 3 dB,表明最小距离法匹配出的结果与受试者的 HRTF 相似。

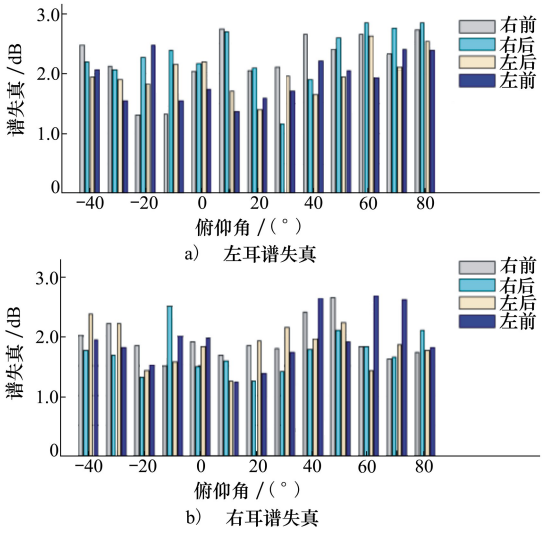


图 6　各俯仰角谱失真

## 4　基于主观听觉实验的匹配算法验证

为进一步验证匹配法的有效性,用含有方位信息的 HRIR 与干信号卷积得到具有方位信息的虚拟声信号,再进行主观方位感知实验。当匹配所得的 HRTF 与受试者的 HRTF 越相似时,虚拟声的方位信息与受试者本身的方位信息越相似,则受试者的主观听觉定位效果应越好。

分别利用 53 人数据库以及人工头 HRTF 合成虚拟声信号。实验时,先测量另选的 20 名受试者的 12 维对 HRTF 影响显著的参数,并利用基于最小距离准则的数据库匹配算法进行匹配,找到数据库中与受试者生理参数最匹配的 HRTF 作为受试者的 HRTF。主观听音测试时受试者分别倾听匹配对象和人工头合成的声信号。

对个人化主观听音结果进行总结整理,个人化和人工头的各方位定位结果如图 7 所示,图 7a),c),e)表示俯仰角 $\phi$ 分别为-40°,0°,40°的个人化 HRTF 定位结果。b),d),f)表示俯仰角 $\phi$ 分别为-40°,0°,40°的人工头 HRTF 定位结果。方块大小表示受试者判断的结果落在该方位的人数占总人数的百分比。斜率为 1 的线表示预测方位与实际方位一致。斜率为-1 的线上则表示前后混淆的结果。

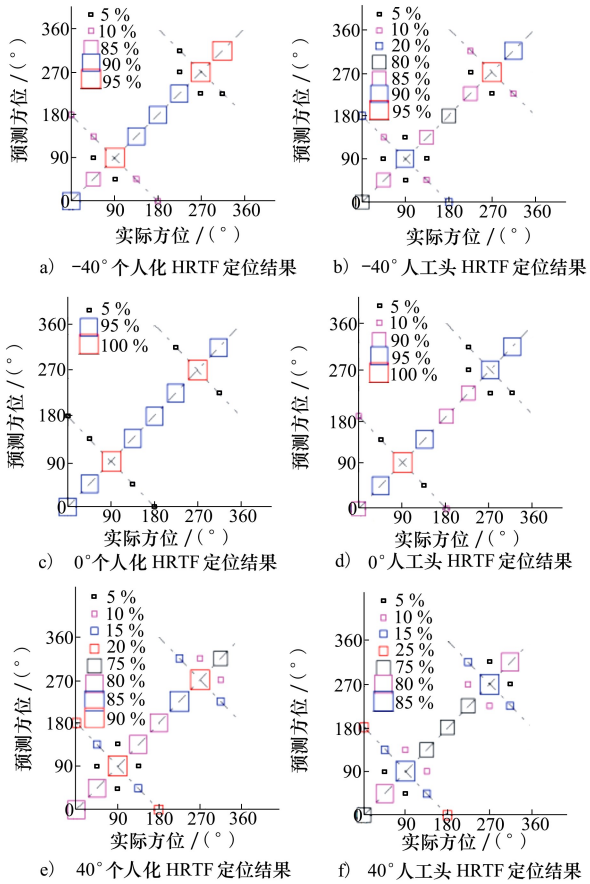


图 7　各俯仰角定位结果

对比个人化和人工头的定位效果,根据各俯仰

角定位错误率和前后混淆率作图 8 和图 9。

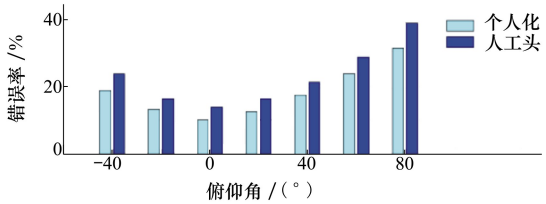


图 8　各俯仰角定位错误率

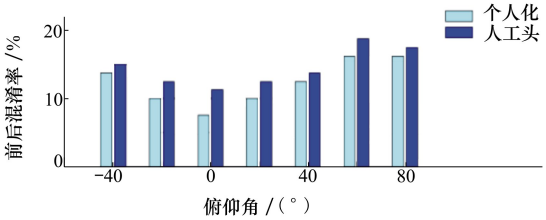


图 9　各俯仰角前后混淆率

对比人工头和个人化 HRTF 对应的主观实验结果发现:

1)基于实测生理参数获得的个人化 HRTF 数据的平均定位错误率和前后混淆率均低于人工头的 HRTF 数据定位结果,说明基于多维生理参数的头相关传递函数个人化方法能够有效地降低前后混淆率,提高定位的精度。

2)测试者对位于左右的声音要比前后的声音感知正确率高,测试者对位于上方的声音要比下方的声音感知正确率高,各俯仰角的定位精度由高到低,依次为 0°,±20°,±40°,60°,80°。

## 5　结　论

本文将主成分分析法和相关分析法相结合,应用于多维生理参数的优选,实现了生理参数的降维,又利用数据库匹配法实现了头相关传递函数的个人化预测,分别通过谱失真和主观听觉实验对预测 HRTF 的有效性进行了验证。结果显示,通过个人化处理,能提高定位准确率 4.6%,降低前后混淆率 2.2% 。

## 参考文献:

[1]　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: 100-101
Xie Bosun. Head Related Transfer Function and Virtual Auditory[M]. Beijing, National Defense Industry Press, 2008: 100-101 (in Chinese)

[2]　曾向阳. 头相关传递函数的个人化方法[J]. 电声技术, 2007, 31(8): 41-46
Zeng Xiangyang. Head-Related Transfer Function Personalization[J]. Electroacoustic Technology, 2007, 31(8): 41-46 (in Chinese)

[3]　Cooper D H. Calculator Program for Head-Related Transfer Functions[J]. Audio Eng Soc, 1982, 30:34-38

[4]　Algazi V R, Duda R O, Duraiswami R, et al. Approximating the Head-Related Transfer Function using Simple Geometric Models of the Head and Torso[J]. Acoust Soc Am, 2002, 112(5): 2053-2064

[5]　Kahana Y, Nelson P A, Petyt M. Boundary Element Simulation of HRTFs and Sound Fields Produced by Virtual Sound Imaging System[C]//AES 105 Convention, San Francisco, CA, USA, 1998, Preprint:4817

[6]　Jin C, Leong P, Leung J, et al. Enabling Individualized Virtual Auditory Space Using Morphological Measurements[C]//Proceedings of the First IEEE Pacific-Rim Conference on Multimedia, 2000, 235-238

[7]　Guillon P, Guignard T, Nicol R. Head-Related Transfer Function Customization by Frequency Scaling and Rotation Shift Based on a New Morphological Matching Method[C]//AES 125th Convention, 2008

[8]　Zotkin D N, Duraiswami R, Grassi E, et al. Fast Head-Related Transfer Function Measurement via Reciprocity[J]. Acoust Soc Am, 2006, 120(4): 2202-2215

[9]　Brown C P, Duda R O. A Structural Model for Binaural Sound Synthesis[J]. IEEE Trans on Speech and Audio Processing, 1998, 6(5): 476-488

[10]　Breebaart J. Effect of Perceptually Irrelevant Variance in Head-Related Transfer Functions on Principal Component Analysis[J]. Journal of the Acoustical Society of America, 2013,133(1):EL1-EL6

[11]　Xie B S. Recoverry of Individual Head-Related Transfer Functions from a Small Set of Measurements[J]. Journal of the Acoustical Society of America, 2012, 132(1): 282-294

[12] 王曙光. 基于数据库匹配的 HRTF 个人化方法[D]. 西安：西北工业大学，2010
Wang Shuguang. HRTF Personalization Method Based on Database Matching[D]. Xi'an, Northwestern Polytechnical University, 2010 (in Chinese)

# Personalization Method for HRTF Based on Multi-Dimensional Physiological Parameters

Huang Wanqiu, Zeng Xiangyang, Wang lei
(School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China)

**Abstract**: Head-Related Transfer Function (HRTF) is the most important factor to achieve Auditory Space Modeling(ASM). HRTF has many applications in the areas of room acoustic modeling, spatial hearing and multimedia. For the reason that HRTF is related to the location and frequency of sound source and the physiological structures of the listener, customization of HRTF becomes the bottle-neck problem of the research and application of ASM. This paper, based on the measured data offers the personalization method of the HRTF based on the multi-dimensional physiological parameters, and the method is verified through the spectral distortion and the subjective auditory localization experiments. The results show that database matching can obtain the personalized HRTF efficiently and in the spatial hearing experiment the personalized HRTF can reduce the front-back confusion of auditory localization and improve the accuracy.

**Keywords**: HRTF; personalization; anthropometrics; correlation analysis; principal component analysis; database matching; auditory location