

◇ 研究报告 ◇

一种 5.1 通路环绕声的动态双耳重放方法*

张承云¹ 谢菠荪^{2†}

(1 广州大学机械与电气工程学院 广州 510006)

(2 华南理工大学物理与光电学院声学研究所 广州 510641)

摘要 为改善 5.1 通路环绕声的双耳重放性能,提出一种基于低价头踪迹跟踪模块的动态双耳重放方法。头踪迹跟踪模块通过单片机采集磁传感器、加速度传感器的输出数据,计算出倾听者头部水平方向信息,并将其经 USB 接口传给计算机进行动态双耳声信号合成。心理声学实验表明,本文提出的方法可以消除虚拟声源前后混乱和头中定位现象,提升 5.1 通路环绕声双耳重放的虚拟声源定位性能。

关键词 头踪迹跟踪模块, 5.1 通路环绕声, 动态, 双耳重放

中图分类号: TN912.27, O249

文献标识码: A

文章编号: 1000-310X(2016)04-0283-05

DOI: 10.11684/j.issn.1000-310X.2016.04.001

A scheme for dynamic binaural reproduction of 5.1 channel surround sound

ZHANG Chengyun¹ XIE Bosun²

(1 School of Mechanical & Electrical Engineering, Guangzhou University, Guangzhou 510006, China)

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

Abstract This work proposes a scheme for improving binaural reproduction of 5.1 channel surround sound by using dynamic rendering with low cost head-tracking module. The head-tracking module consists of a magnetic sensor and an accelerometer. A microcontroller unit is used to calculate the horizontal orientation of the listener's head. The orientation data is transmitted to the PC via the serial port for dynamic binaural synthesis. Psychoacoustic experiment indicates that the scheme eliminates the front-back confusion and in-head-localization and therefore improves the localization performance in the binaural 5.1 channel reproduction.

Key words Head-tracking module, 5.1 channel surround sound, Dynamic, Binaural reproduction

2016-06-01 收稿; 2016-06-14 定稿

*广州市科技计划项目 (2014-Y2-00021), 广州市属高校科研项目 (1201411212)

作者简介: 张承云 (1975-), 男, 江西于都人, 博士, 副教授, 研究方向: 声频信号处理。

†通讯作者 E-mail: phbsxie@scut.edu.cn

1 引言

5.1 通路环绕声原本是为扬声器重放而设计的。针对不同的应用, 5.1 通路环绕声有时也需要通过耳机来重放, 但这样会使得声音空间信息受损, 不能获得正确的立体声声像, 产生头中定位效应。为改善耳机重放 5.1 通路环绕声的效果, 可采用各扬声器方向的头相关传输函数 (Head related transfer functions, HRTF) 对信号进行滤波, 合成双耳声信号馈给耳机进行重放^[1]。这相当于用信号处理的方法虚拟出各个通路扬声器。这类耳机虚拟重放技术虽然可以产生一定的声音空间信息, 但还是存在一定缺陷, 例如虚拟声源前后混乱、部分方向仍存在头中定位现象。通过采用个性化 HRTF 进行信号处理以及进行个性化的耳机-外耳传输特性均衡, 可以提高合成双耳信号的精确度^[2], 但也只能部分改善上述缺陷。

现实中倾听者头部是会运动的, 会带来动态听觉信息, 这对主观感知非常重要^[3]。现有研究表明, 包含了动态信息处理的耳机虚拟听觉实时绘制可以明显降低虚拟声源前后混乱率、提高定位准确度、消除头中定位效应并使感知的虚拟声源距离更远^[4-7]。动态信息是通过头踪迹跟踪器检测倾听者头部运动来获取, 包括 6 个自由度 (3 个平移自由度和 3 个转动自由度)。若在 5.1 通路环绕声双耳重放时加入动态信息, 预计能有效的改善听觉效果。5.1 通路环绕声技术一般用于消费类电子产品, 然而现有的头踪迹跟踪器产品一般体积大并且价格较高, 不适合集成到普通消费类电子产品中。由于 5.1 通路环绕声只考虑产生平面上的声音空间信息, 因此可以只处理水平方向这一个动态信息。这样对头踪迹跟踪器的性能要求可以降低, 只要能输出水平角一个自由度参数即可。但要求头踪迹跟踪器体积小、重量轻、价格低, 对便携式产品还要求功耗小。

本文采用磁传感器和加速度传感器设计一种低价的水平角度检测模块, 用于跟踪头部转动时的水平角变化。通过 USB 接口将此数据传给 PC 机, 在 PC 机上用 C++ 语言实现动态的 5.1 通路环绕声双耳重放。

2 5.1 通路环绕声的动态双耳重放原理

实现动态的 5.1 通路环绕声耳机虚拟重放系统结构如图 1 所示。

用扬声器重放 5.1 通路环绕声时, 共有前方左 (L)、右 (R)、中 (C) 和左环绕 (LS)、右环绕 (RS) 五个独立的全频带通路, 加上一个低频效果通路 LFE 。按国际电信联盟 (ITU) 所推荐的标准^[8], 五个全频带扬声器的方位角为

$$\begin{aligned}\theta_L &= 330^\circ, \quad \theta_R = 30^\circ, \quad \theta_C = 0^\circ, \\ \theta_{LS} &= 250^\circ \pm 10^\circ, \quad \theta_{RS} = 110^\circ \pm 10^\circ,\end{aligned}\quad (1)$$

其中 $\theta = 0^\circ$ 为正前方, $\theta = 90^\circ$ 为正右方。

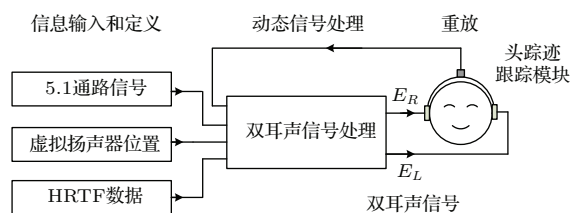


图1 5.1 通路环绕声的动态双耳重放系统结构

Fig. 1 Blocked diagram of dynamic binaural reproduction of 5.1 channel surround sound

设前方左、中、右扬声器到倾听者双耳的 HRTF 分别为 H_{LL} 、 H_{RL} 、 H_{LC} 、 H_{RC} 、 H_{LR} 、 H_{RR} , 左、右环绕扬声器到双耳的 HRTF 分别为 H_{LLS} 、 H_{RLS} 、 H_{LRS} 、 H_{RRS} 。由于低频效果通路 LFE 的处理与中置通路 C 类似, 所以在以下分析中略去。采用耳机重放时可采用 (2) 式合成双耳声信号^[1]:

$$\begin{aligned}E_L &= H_{LL}L + H_{LC}C + H_{LR}R \\ &\quad + H_{LLS}LS + H_{LRS}RS, \\ E_R &= H_{RL}L + H_{RC}C + H_{RR}R \\ &\quad + H_{RLS}LS + H_{RRS}RS,\end{aligned}\quad (2)$$

这样重放的双耳声压将等于或正比于扬声器重放的情况, 产生正确的声音空间信息。

对于稳态情况, (2) 式中的各个 HRTF 是不变的。但对于动态情况, 由于各个虚拟扬声器在空间的绝对位置不变, 当倾听者头部转动后, 虚拟扬声器与倾听者头部的相对位置会改变, 双耳声信号需动态合成^[9-10]。这里只考虑水平角这一个动态信息, 若某虚拟扬声器的初始水平角坐标为 θ_0 , 倾听者头部水平转动 $\Delta\theta$ 后, 则虚拟扬声器新的水平角坐标为

$$\theta = \theta_0 - \Delta\theta, \quad (3)$$

其中 $\Delta\theta$ 由头踪迹跟踪器模块测量得到。信号处理时需根据虚拟扬声器新的水平角坐标 θ 不断刷新 (2) 式的各个 HRTF 数据, 从而实现动态信号处理。

HRTF是声源空间位置的连续函数,但实验测量得到的是有限离散空间方向的HRTF数据。为了虚拟水平面相对倾听者头部任意角度位置的声源(扬声器),需要通过空间插值的方法得到未测量方向的HRTF数据。现有的多种不同空间插值方法都可以用于HRTF的方向插值,比如线性插值的方法,采用相邻两个测量方向的HRTF的线性组合来近似未测量方向的HRTF。利用(3)式的关系计算出虚拟扬声器相对倾听者新的头部位置参数后,插值得到对应的HRTF数据,代入(2)式即可实现动态信号处理。由于实际处理中,只能以一定的时间间隔对倾听者位置进行检测(时间采样),直接从一个位置的HRTF切换到另一个位置的HRTF容易产生可察觉切换噪声。这可以通过对HRTF切换前后的输出信号进行过渡处理,或对HRTF进行主成分分解(PCA),HRTF切换时改变相关的权重系数^[11-12]。其中PCA处理办法更简单高效,并且不会产生可察觉切换噪声^[1]。

3 动态双耳重放系统软硬件设计

3.1 硬件设计

硬件部分包括一台个人计算机(Intel i3 CPU@2.53GHz, 2G内存)、支持ASIO功能的声卡(ESI UGM96)、耳机(Sennheiser IE8)、头踪迹跟踪模块。计算机是整个系统的信息处理中心,系统开始工作前计算机先通过USB接口下传参数完成头踪迹跟踪模块及声卡的初始化工作,以及设置输入信号、虚拟扬声器位置和HRTF数据等信息。系统开始工作后头踪迹跟踪模块将倾听者头部的水平方向信息通过USB接口传给计算机,用于信号处理软件合成双耳声信号,然后通过声卡传送给耳机重放。

头踪迹跟踪模块主要由磁传感器、加速度传感器、单片机、串口转USB芯片组成,如图2所示。建立如图3所示的坐标系, $O-xyz$ 表示参考坐标系,其中 xOy 表示水平面, x 轴指向北极。 $O_1-x_1y_1z_1$ 表示磁传感器的坐标系, x_1 轴在水平面的投影与 x 轴之间的角度定义为水平角 α ,顺时针为正。磁传感器可以测出在此坐标系中的三个地磁场分量 H_{x1} 、 H_{y1} 、 H_{z1} 。 $O_2-x_2y_2z_2$ 表示加速度传感器的坐标系,加速度传感器可以测出重力加速度 g 在此坐标系中的三个轴分量 g_{x2} 、 g_{y2} 、 g_{z2} 。这两个传感器测出的数据都传给单片机进行处理。正式读数之前磁传感器和加

速度传感器都需先校准。

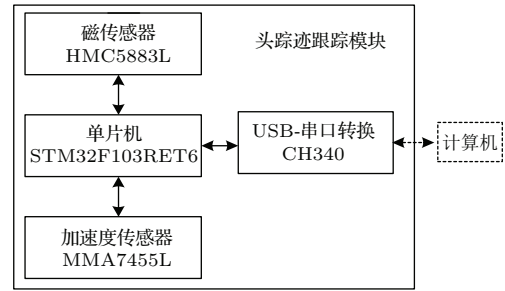


图2 头踪迹跟踪器模块

Fig. 2 Blocked diagram of head tracking module

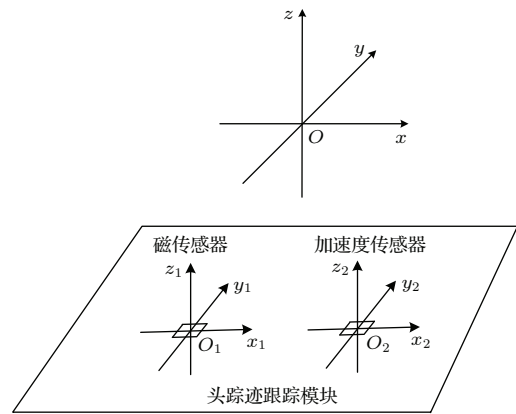


图3 坐标系统

Fig. 3 Coordinate system

当头踪迹跟踪器模块水平放置时,单片机先根据 H_{x1} 、 H_{y1} 地磁分量就可以利用(4)式计算出角度 α ,再根据 H_{x1} 、 H_{y1} 的正负号将 α 调整到 $0^\circ \leq \alpha < 360^\circ$ 。但一般情况下,模块并不是水平放置,两个传感器的坐标系发生旋转。将传感器坐标系绕参考坐标系 x 轴旋转的角度称为倾斜角 β ,绕 y 轴旋转的角度称为滚动角 γ 。这时直接根据 H_{x1} 、 H_{y1} 计算出的角度 α 会有误差,需要先利用 g_{x2} 、 g_{y2} 、 g_{z2} 根据(4)式计算出倾斜角 β 与滚动角 γ ,然后利用(5)式对 H_{x1} 、 H_{y1} 进行修正,最后再利用(4)式计算出角度 α 。单片机将 α 的数据通过串口传给计算机进行信号处理,前后两个 α 的差值即为(3)式的 $\Delta\theta$ 。

$$\begin{aligned}\alpha &= \arctan\left(\frac{H_{y1}}{H_{x1}}\right), \\ \beta &= \arcsin\left(\frac{g_{y2}}{g}\right), \\ \gamma &= \arcsin\left(\frac{g_{x2}}{g}\right),\end{aligned}\quad (4)$$

$$H'_{x1} = H_{y1} \sin \beta \sin \gamma + H_{z1} \cos \beta \sin \gamma,$$
$$H'_{y1} = H_{y1} \cos \beta - H_{z1} \sin \beta.$$

(5)

地磁场的强度和方向与所测位置的经度、纬度、高度有关,因此磁传感器测量的地磁场数据会有所不同,另外测量结果还会受周边环境各种材料形成的磁场影响。数据处理时要对磁传感器的测量结果进行校正补偿^[13-14]。

3.2 软件设计

如图4所示,软件分成两部分:计算机上的信号处理软件和头踪迹跟踪模块软件。信号处理软件在Microsoft Visual Studio.NET 2010平台上用C++语言通过多线程实现,包括设定虚拟扬声器位置、声频数据输入输出、头踪迹跟踪数据接收、信号处理参数计算、信号滤波及合成等功能。头踪迹跟踪模块软件在Keil uVision4平台上用C语言实现,包括磁传感器、加速度传感器、单片机及串口初始化,磁传感器与加速度传感器三轴数据读取、计算角度 α 、串口数据收发等功能,采用单线程处理。

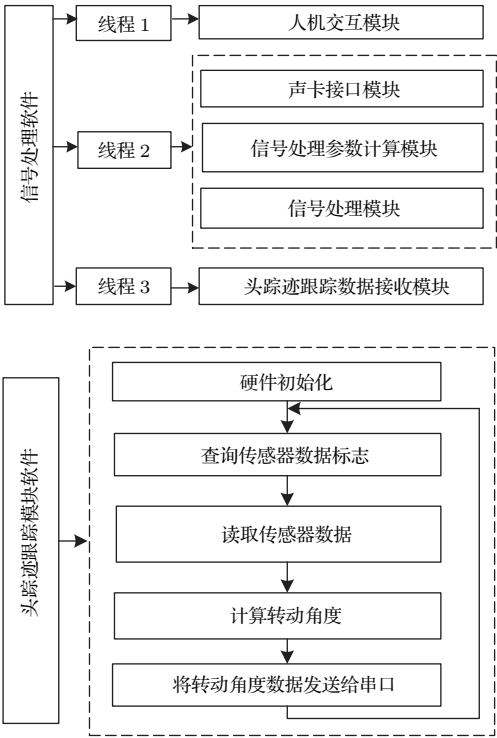


图4 软件结构图

Fig. 4 The software structure

4 系统客观性能参数

场景刷新率、系统滞后时间是虚拟听觉动态处理的重要客观性能参数^[15-16]。

场景刷新率指的是在单位时间内信号处理(场景)的刷新次数。它主要取决于头踪迹跟踪数据刷新率。磁传感器HMC5883的数据刷新率为160 Hz、加速度传感器MMA7455L的数据刷新率为125 Hz。利用计算机内部的高精度定时器,测量出头踪迹跟踪模块输出数据的时间间隔平均值为8.05 ms,折算成刷新率约为124 Hz。

系统滞后时间是指倾听者头部运动到某个位置到双耳听到相应信号的时间间隔。它受头踪迹跟踪模块、声卡、信号处理、数据通信、数据缓存的影响,其中前两项影响最大。UGM96声卡的ASIO数据缓冲区取128点,带来的延时为2.9 ms(采样率为44.1 kHz)。头踪迹跟踪模块的延时主要由磁传感器HMC5883、加速度传感器MMA7455L及数据处理部分的延时组成。两种传感器的具体延时值未知,但不会超过20 ms,数据处理部分的延时小于8 ms,串口发送数据的延时小于0.1 ms。因此系统滞后时间能控制在40 ms之内。

表1为系统的性能参数。已有研究表明,场景刷新率大于20 Hz,系统滞后时间小于60 ms,在动态双耳声信号重放中能得到令人满意的动态性能^[15-16]。而本系统的动态参数都在这个范围之内。

表1 系统性能参数

Table 1 Objective performance of the system

系统滞后时间(ms)	系统刷新率(Hz)	滤波器长度(taps)
40	124	128

5 心理声学实验

心理声学实验的目的是验证动态双耳重放的效果。实验采用全可听频带白噪声作为原始信号,白噪声信号在计算机中通过软件产生(采样率44.1 kHz、量化精度16 bit、长度5 s)。采用MIT媒体实验室测量的KEMAR人工头的远场HRTF数据^[17],分别作稳态和动态自由场虚拟声源信号处理。

选择水平面上5个虚拟声源方向,方位角 $\theta = 0^\circ, 30^\circ, 110^\circ, 250^\circ, 330^\circ$ 。共有6名受试者(男女各3名)参加实验。对稳态和动态信号处理,受试者对虚拟声源的水平角(θ)、仰角(ϕ)和距离(头内或头外)进行判断^[18]。每个虚拟声源由每名受试者按随机顺序作3次判断,6名受试者共有18次判断。

表2 虚拟声源定位实验数据统计结果
Table 2 Statistical results for virtual source localization experiment

虚拟扬声器位置	稳态信号处理				动态信号处理			
	头内定位率 (%)	前后混乱率 (%)	$\Delta\phi(^{\circ})$	$\Delta\theta(^{\circ})$	头内定位率 (%)	前后混乱率 (%)	$\Delta\phi(^{\circ})$	$\Delta\theta(^{\circ})$
前方	18.5	18.2	11.2	14.3	0	0	8.6	8.4
环绕	0	25	7.8	9.9	0	0	7.7	9.5

实验结果按前方虚拟声源方向和环绕虚拟声源方向分别统计出头内定位率、前后混乱率,对镜像方向混乱的数据进行空间反演后再统计出水平角平均角度误差($\Delta\theta$)、仰角平均角度误差($\Delta\phi$)。表2给出了数据统计结果。数理统计的T检验($\alpha = 0.05$)结果表明,动态信号处理可以消除头中定位效应,对前方虚拟声源头外化有明显作用;能消除虚拟声源前后混乱现象;减小前方虚拟声源的水平角平均角度误差、仰角平均角度误差。

6 结论

设计了一种低价头踪迹跟踪模块,用于基于个人计算机的5.1通路环绕声的动态双耳重放。心理声学实验结果表明,这种动态双耳重放方法能消除稳态双耳重放中常出现的虚拟声源前后混乱及头中定位现象。因此本文提出的方法可以改善5.1通路环绕声双耳重放的虚拟声源定位效果。

参 考 文 献

- [1] XIE B S. Head related transfer function and virtual auditory display[M]. 2nd Edition. Fort Lauderdale: J Ross Publishing, 2013.
- [2] XIE B S, ZHANG C Y, ZHONG X L. A cluster and subjective selection-based HRTF customization scheme for improving binaural reproduction of 5.1 channel surround sound[C]. AES 134th Convention, 2013.
- [3] BLAUERT J. Spatial hearing [M]. Revised Edition. Cambridge: MIT Press, 1997.
- [4] MILLER J D, WENZEL E M. Recent developments in SLAB: A software-based system for interactive spatial sound synthesis[C]. International Conference on Auditory Display, 2002.
- [5] SAVIOJIA L, HUOPANIEMI J, LOKKI T, et al. Creating interactive virtual acoustic environments[J]. J. Audio Eng. Soc., 1999, 47(9): 675-705.
- [6] BLAUERT J, LEHNERT H, SAHRHAGE J, et al. An interactive virtual-environment generator for psychoacoustic research. I: Architecture and implementation[J]. Acta Acustica United with Acustica, 2000, 86(1): 94-102.
- [7] ZHANG C Y, XIE B S. Platform for virtual auditory environment real time rendering system[J]. Chin. Sci. Bull., 2013, 58(3): 316-327.
- [8] ITU. ITU-R BS 775-2, Multichannel stereophonic sound system with and without accompanying picture[S].
- [9] WENZEL E M, ARRUDA M, KISTLER D J, et al. Localization using nonindividualized head-related transfer functions[J]. J. Acoust. Soc. Am., 1993, 94(1): 111-123.
- [10] WIGHTMAN F L, KISTLER D J. Resolution of front-back ambiguity in spatial hearing by listener and source movement[J]. J. Acoust. Soc. Am., 1999, 105(5): 2841-2853.
- [11] LARCHER V, JOT J M, GUYARD J, et al. Study and comparison of efficient methods for 3D audio spatialization based on linear decomposition of HRTF data[C]. AES 108th Convention, 2000.
- [12] XIE B S, ZHANG C Y. An algorithm for efficiently synthesizing multiple near-field virtual sources in dynamic virtual auditory display[C]. AES 132nd Convention, 2012.
- [13] 龙礼, 张合. 三轴地磁传感器误差的自适应校正方法[J]. 仪器仪表学报, 2013, 34(1): 161-165.
LONG Li, ZHANG He. Automatic and adaptive calibration method of tri-axial magnetometer[J]. Chin. J. Sci. Instrum., 2013, 34(1): 161-165.
- [14] 龙达峰, 刘俊, 李杰, 等. 地磁传感器误差参数估计与补偿方法[J]. 传感技术学报, 2014, 27(12): 1649-1653.
LONG Dafeng, LIU Jun, LI Jie, et al. Error parameters estimation and compensation method apply to magnetometer[J]. Chin. J. Sens. Actuators, 2014, 27(12): 1649-1653.
- [15] BRUNGART D S, KORDIK A J, SIMPSON B D. Effects of headtracker latency in virtual audio displays[J]. J. Audio Eng. Soc., 2006, 54(1/2): 32-44.
- [16] SANDVAD J. Dynamic aspects of auditory virtual environments[C]. AES 100th Convention, 1996.
- [17] GARDNER W G, MARTIN K D. HRTF measurements of a KEMAR[J]. J. Acoust. Soc. Am., 1995, 97(6): 3907-3908.
- [18] PERNAUX J M, EMERIT M, NICOL R. Perceptual evaluation of binaural sound synthesis: the problem of reporting localization judgements[C]. AES 114th Convention, 2003.