



(12)发明专利

(10)授权公告号 CN 105509691 B

(45)授权公告日 2018.01.26

(21)申请号 201510876214.X

(51)Int.Cl.

(22)申请日 2015.12.03

G01B 21/22(2006.01)

(65)同一申请的已公布的文献号

G01P 15/18(2013.01)

申请公布号 CN 105509691 A

G01C 19/00(2013.01)

(43)申请公布日 2016.04.20

审查员 乐兵兵

(66)本国优先权数据

201510738104.7 2015.11.03 CN

(73)专利权人 北京时代拓灵科技有限公司

地址 100085 北京市海淀区上地信息路26号01层0014-32室

(72)发明人 孙学京 刘皓

(74)专利代理机构 北京创遇知识产权代理有限公司 11577

代理人 吕学文 武媛

权利要求书2页 说明书10页

(54)发明名称

多传感器组融合的检测方法及支持头部跟踪的环绕声方法

(57)摘要

本发明涉及一种多传感器组融合的运动检测方法和一种支持头部跟踪的虚拟环绕声生成方法,所述运动检测方法采用多个传感器组采集运动相关信息,依据各单组检测数据计算获得最终检测数据,所述虚拟环绕声生成方法基于主要由耳机和头戴式显示器组成的虚拟现实设备,所述头戴式显示装置上设有第一传感器组,所述耳机上设有第二传感器组,通过上述运动检测方法检测用户头部运动角度,依据用户头部运动角度对音频进行旋转,生成能够跟踪头部运动的虚拟环绕声并通过耳机播放。采用本发明的检测方法能够获得更精确和更可靠的运行检测结果,采用本发明的环绕声生成方法可以跟踪用户的头部运动,获得高质量的听觉效果。

1. 一种多传感器组融合的运动检测方法,用于检测运动数据,采用多个传感器组采集运动相关信息,对各传感器组进行传感器融合,获得该传感器组的单组检测数据,依据各单组检测数据计算获得最终检测数据。

2. 如权利要求1所述的多传感器组融合的运动检测方法,其特征在于所述依据各单组检测数据计算获得最终检测数据的方法为:计算能够有效工作的各传感器组的单组检测数据的算术平均值或加权平均值,以获得的算术平均值或加权平均值为所述的最终检测数据,当只有一组传感器组能够有效工作时,以该传感器组的单组检测数据为最终检测数据。

3. 如权利要求2所述的多传感器组融合的运动检测方法,其特征在于在计算加权平均值时,依据相关各传感器组的精确度、可靠性和/或关联性确定各传感器组对应的权重系数,精确度高的传感器组的权重系数高于精确度低的传感器组的权重系数,可靠性高的传感器组的权重系数高于可靠性低的传感器组的权重系数,关联性高的传感器组的权重系数高于关联性低的传感器组的权重系数。

4. 如权利要求1、2或3所述的多传感器组融合的运动检测方法,其特征在于所述传感器组中包括加速度计、陀螺仪和磁力计,所述运动数据为角度,所述单组检测数据为单组检测角度,所述最终检测数据为最终检测角度,所述传感器融合方式为依据下列公式计算单组检测角度:

$$\theta_{f,i}(t) = \alpha_{1,i}(\theta_{f,i}(t-1) + \omega_{g,i}(t) dT) + (1 - \alpha_{1,i}) \theta_{am,i}(t)$$

其中, $\theta_{f,i}(t)$ 为第*i*传感器组在*t*时刻的单组检测角度, $\theta_{f,i}(t-1)$ 为第*i*组传感器组在*t*时刻之前一个采样间隔时的单组检测角度;

$\alpha_{1,i}$ 为第*i*传感器组的时间常数;

$\omega_{g,i}(t)$ 为依据第*i*传感器组中的陀螺仪获得的角速度;

dT_i 为第*i*传感器组的采样时间间隔,各传感器组的采样间隔时间相同或不相同;

$\theta_{am,i}(t)$ 为依据第*i*传感器组中的加速度计和磁力计获得的*t*时刻角度;

*i*为传感器组序号,为从1到*N*的自然数,其中*N*为传感器组的数量,大于等于2。

5. 如权利要求4所述的多传感器组融合的运动检测方法,其特征在于所述传感器组中加速度计、陀螺仪和磁力计的数量均为一个,且分别为3-轴加速度计、3-轴陀螺仪和3-轴磁力传感器。

6. 如权利要求5所述的多传感器组融合的运动检测方法,其特征在于所述传感器组的数量为两组,所述两传感器组的采样时间间隔相同,时间常数不同。

7. 一种支持头部跟踪的虚拟环绕声生成方法,基于主要由耳机和头戴式显示器组成的虚拟现实设备,其特征在于采用权利要求1-6中任意一项权利要求所述的方法检测用户头部运动角度,依据用户头部运动角度对音频进行旋转,消除因用户头部运动带来的方向失真,生成能够跟踪头部运动的虚拟环绕声并通过耳机播放,所述传感器组的数量为两组,包括第一传感器组和第二传感器组,所述第一传感器组设置在所述头戴式显示装置上,所述第二传感器组设置在耳机上。

8. 如权利要求7所述的虚拟环绕声生成方法,其特征在于所述头戴式显示器为宿主单元,存有和/或能够从外部获得音频和视频内容,所述头戴式显示器设有主数据处理单元,在头戴式显示器能够正常工作的情况下,用于依据第一传感器组和第二传感器组的单组检测角度计算出用户头部运动角度的最终检测数据,依据用户头部运动角度对音频信号进行

旋转并生成所述能够跟踪头部运动的虚拟环绕声。

9. 如权利要求8所述的虚拟环绕声生成方法,其特征在於当所述头戴式显示器的电量不足时,直接传输音频到耳机中,并采用下列任意一种方式处理:

方式一:由所述耳机的处理单元依据第一传感器组和第二传感器组的单组检测角度计算获得用户头部运动角度的最终检测数据,并依据所获得的用户头部运动角度进行音频信号的旋转,生成所述跟踪头部运动的虚拟环绕声;

方式二:以第二传感器组的单组检测角度用户头部运动角度的最终检测数据,由所述耳机的处理单元进行音频信号的旋转,生成所述跟踪头部运动的虚拟环绕声,不关闭第一传感器组;

方式三:以第二传感器组的单组检测角度用户头部运动角度的最终检测数据,由所述耳机的处理单元进行音频信号的旋转,生成所述跟踪头部运动的虚拟环绕声,关闭第一传感器组或者关闭涉及第一传感器组的融合算法模块,并将第二传感器组的单组检测数据传到所述头戴式显示器。

10. 如权利要求9所述的虚拟环绕声生成方法,其特征在於所述第一传感器组和第二传感器组在各自传感器融合计算中所用的时间常数不相同,其中一个传感器组的时间常数为0.95,另一个传感器组的时间常数为0.99。

多传感器组融合的检测方法及支持头部跟踪的环绕声方法

技术领域

[0001] 本发明涉及一种多传感器组融合的运动检测方法,还涉及一种基于这种多传感器组融合方法的支持头部跟踪的虚拟环绕声生成方法。

背景技术

[0002] 用户戴着耳机听虚拟环绕声时,当用户头部旋转时,耳机里的虚拟环绕声会跟着用户的头部旋转,这样导致人在现场听音乐的感觉不同,生成的虚拟环绕声不够真实,因此,需要使用运动传感器进行头部运动跟踪(head tracking),使虚拟现实设备或头戴式显示器(head-mounted display,HMD)的使用者(用户)得到不受头部运动影响的优质环绕声。

[0003] 现有运动传感器主要包括加速度计、陀螺仪和磁力传感器,这些传感器在运动跟踪和绝对方向方面都有自己各自固有的强项和弱点。例如,加速度计提供一个重力向量(指向地球中心的向量),而磁力计则是一个指南针,这两种传感器的信息可以用来计算设备的方向,然而这两个传感器的输出并不精确,包含了大量的噪音,而陀螺仪提供沿着三个轴旋转的角速度,这个信息非常准确,并且反应很快,但长时间会产生漂移误差,其原因在于需要对角速度进行积分来得到方向信息,而积分过程会导致微小数值误差,误差长时间的积累就形成了比较明显的漂移。为扬长避短,将上述三种传感器组合成传感器组,对全部传感器的信号依据适宜的方式进行运算,形成更加精确的运动检测结果,就是传感器的融合(sensor fusion),人们已经进行过传感器融合的研究并取得了一些有效的运算方式,但进一步改进和优化融合运算依然有助于技术进步,特别是,单一传感器组的误差相对较大甚至会出现一定的故障,如何获得更为精确的检测结果并保证在局部故障时依然能够有效运行,就成为目前需要解决的技术问题。

发明内容

[0004] 为了克服现有技术的上述缺陷,本发明提供了一种多传感器组融合的运动检测方法,还提供了一种基于这种运动检测方法的支持头部跟踪的环绕声生成方法,采用这种多传感器组融合方法能够获得更精确和更可靠的运行检测结果,采用这种环绕声生成方法可以跟踪用户的头部运动,获得高质量的听觉效果。

[0005] 本发明的技术方案为:

[0006] 一种多传感器组融合的运动检测方法,用于检测运动数据,采用多个传感器组采集运动相关信息,对各传感器组进行传感器融合,获得该传感器组的单组检测数据,依据各单组检测数据计算获得最终检测数据。

[0007] 一种支持头部跟踪的虚拟环绕声生成方法,基于主要由耳机和头戴式显示器组成的虚拟现实设备,采用本发明公开的任意一种运动检测方法检测用户头部运动角度,依据用户头部运动角度对音频进行旋转,消除因用户头部运动带来的方向失真,生成能够跟踪头部运动的虚拟环绕声并通过耳机播放,所述传感器组的数量为两组,包括第一传感器组和第二传感器组,所述第一传感器组设置在所述头戴式显示装置上,所述第二传感器组设

置在耳机上。

[0008] 本发明的有益效果是：通过传感器组内各传感器融合，避免了单一类型传感器的不足，自各传感器的信号组合在一起，产生一个更加精确的单组检测数据；通过采用多组传感器并依据多组传感器计算最终检测数据，进一步提高劳动数据的精确性和可靠性，且保证了在部分传感器或传感器组出现故障时，系统依然能够正常运动；由于依据头部运动角度生成相应的环绕声，避免了因头部运动导致的方向性失真，明显地改善了用户体验；由于将两组传感器分别设置头戴式显示器和耳机上，不仅避免带来单一传感器组可能存在的误差和故障，提高了对用户头部运动角度的检测精度，提高了环绕声的质量，而且在不增加用户负担或不适感的情况下，有效地检测了头戴式显示器和耳机两方面的运动角度，减小或避免了偶然因素导致头戴式显示器或耳机移动所产生的音频旋转；由于两组传感器在传感器融合中可以不同的时间常数，有助于兼顾两种时间常数的优势，进一步改善音质；由于可以最终检测数据计算中可以对不同的单组检测数据设置不同的权重，有助于进一步提高检测结果的精确性和可靠性；由于在头戴式显示器电量不足时可以充分利用耳机的处理单元承担相关数据处理，节省能耗。

具体实施方式

[0009] 本发明提供的多传感器组融合的运动检测方法，可用于检测运动数据，其采用多个传感器组采集运动相关信息，对各传感器组进行传感器融合，获得该传感器组的单组检测数据，依据各单组检测数据计算获得最终检测数据。这种检测方法可以用于虚拟现实播放设备的用户头部跟踪，以便更为精确地获得用户头部旋转角度，依据用户头部旋转角度调整所播放的音频、视频和游戏画面等，以消除因头部旋转导致的失真。这种检测方法也可以用于检测其他运动信号和/或用于其他场合，例如检测速度、距离及其他任意适应于多传感器组检测的数据。

[0010] 优选的，所述依据各单组检测数据计算获得最终检测数据的方法为：计算能够有效工作的各传感器组的单组检测数据的算术平均值或加权平均值，以获得的算术平均值或加权平均值为所述的最终检测数据，当只有一组传感器组能够有效工作时，以该传感器组的单组检测数据为最终检测数据。

[0011] 优选的，在计算加权平均值时，依据相关各传感器组的精确度、可靠性和/或关联性确定各传感器组对应的权重系数，精确度高的传感器组的权重系数高于精确度低的传感器组的权重系数，可靠性高的传感器组的权重系数高于可靠性低的传感器组的权重系数，关联性高的传感器组的权重系数高于关联性低的传感器组的权重系数，所述精确度可以依据所用传感器本身的精确度确定，所述可靠性可以依据实验分析结果确定或依据传感器本申请的可靠性能确定，所述关联性可以依据排除传感器因素后的系统误差确定，依据实验或对系统构造的分析获得相关系统误差，系统误差大的关联性低，由此增加优质传感器组数据在最终检测结果所占的比重，以提高检测精度。

[0012] 所述传感器组中可以包括加速度计、陀螺仪和磁力计(磁力传感器)。以充分利用加速度计和磁力计在方向检测上没有积累误差的优势，避免其带有大量噪音的劣势，充分利用陀螺仪在角速度检测中信息非常准确且并且反应很快的优势，避免其长时间会产生漂移误差的劣势。

[0013] 对于虚拟现实播放等场合,将所述运动数据设定为角度,相应地,所述单组检测数据为单组检测角度,所述最终检测数据为最终检测角度。

[0014] 所述传感器融合方式可以为依据下列公式计算单组检测角度:

$$[0015] \quad \theta_{f,i}(t) = \alpha_{1,i}(\theta_{f,i}(t-1) + \omega_{g,i}(t) dT) + (1 - \alpha_{1,i}) \theta_{am,i}(t)$$

[0016] 其中, $\theta_{f,i}(t)$ 为第*i*传感器组在*t*时刻的单组检测角度, $\theta_{f,i}(t-1)$ 为第*i*组传感器组在*t*时刻之前一个采样间隔时(即*t-dT*时刻)的单组检测角度,也可以表示为 $\theta_{f,i}(t-dT)$;

[0017] $\alpha_{1,i}$ 为第*i*传感器组的时间常数,体现了 $\theta_{f,i}(t)$ 的平滑程度,可以依据实验和/或需要等因素设定;

[0018] $\omega_{g,i}(t)$ 为依据第*i*传感器组中的陀螺仪获得的角速度;

[0019] dT_i 为第*i*传感器组的采样时间间隔,各传感器组的采样间隔时间可以相同,也可以不同;

[0020] $\theta_{am,i}(t)$ 为依据第*i*传感器组中的加速度计和磁力计获得的*t*时刻角度;

[0021] *i*为传感器组序号,为从1到*N*的自然数,其中*N*为传感器组的数量,大于等于2。

[0022] 所述传感器组中加速度计、陀螺仪和磁力计的数量可以均为一个,且分别为3-轴加速度计、3-轴陀螺仪和3-轴磁力传感器。

[0023] 通过上述计算公式,可以使用短时的陀螺仪信息和长时的加速度计和磁力计信息,这就相当于对陀螺仪输出使用了一个高通滤波器,而对其他传感器输出使用了低通滤波器,系数 $\alpha_{1,i}$ 是时间常量,用来控制低通滤波器的截止频率,也就是 $\theta_{am,i}(t)$ 的平滑程度。时间常量 $\alpha_{1,i}$ 的选取对系统性能有重大的影响,数值过大则造成系统更多依赖于陀螺仪角度,对加速度计的角度响应缓慢,导致系统漂移错误加大,数值过小则造成角度平滑不够,包含大量加速度计结果噪音,导致准确性下降。 $\alpha_{1,i}$ 值的选取依赖于多个因素,如传感器本身的特性,可以根据经验人为选取一组数值,但通常数值是一种折中的选法。

[0024] 所述传感器组的数量可以为两组,所述两传感器组的采样时间间隔可以相同,时间常数可以不同。由于传感器准确性受多种因素制约。由于单独传感器往往不能达到要求,当用于虚拟现实头戴现实设备和虚拟现实耳机上上时,采用两组传感器组来解决上述问题,可以提高整体的准确性。在进行两组传感器单独检测数据的加权平均值时,权重系数*w*可以事先人为确定,但需满足 $w_1 + w_2 = 1$ 。如果已知某一组传感器精准度会高些(如使用了款高端耳机加低端手机),则可以把其相应的权重加大,甚至是 $w_1 = 1$ 。权重系数的设定还可基于在线统计分析,即对两组传感器输出结果的可靠性进行分析从而导出权重系数。例如我们可以计算 $\theta_{f,1}(t)$ 和 $\theta_{f,2}(t)$ 的方差,对方差较小的传感器增加权重,现有技术下还存在其他若干相关技术方法,也可以用于本发明中。

[0025] 一种支持头部跟踪的虚拟环绕声生成方法,基于主要由耳机和头戴式显示器组成的虚拟现实设备,采用本发明公开的任意一种运动检测方法检测用户头部运动角度,依据用户头部运动角度对音频进行旋转,消除因用户头部运动带来的方向失真,生成能够跟踪头部运动的虚拟环绕声并通过耳机播放,所述传感器组的数量为两组,包括第一传感器组和第二传感器组,所述第一传感器组设置在所述头戴式显示装置上,所述第二传感器组设置在耳机上。所述耳机和头戴式显示器之间可以采用蓝牙或其他任意适宜的方式通信,传送相关数据。根据实际需要,也可以在头戴式显示器和/或耳机上设置更多个传感器组,下面以在头戴式显示器和耳机上各设置一个传感器组为例说明相应的环绕声生成方法,当采

用更多个传感器组时,头部旋转角度的计算依据本发明提供的多传感器组融合的运动检测方法,获得头部旋转角度后对音频的处理与在头戴式显示器和耳机上各设置一个传感器组时相同。

[0026] 所述音频的内容可以是ambisonic声场,或者是object audio等任意适宜的形式。

[0027] 优选的,所述头戴式显示器为宿主单元,存有和/或能够从外部获得音频和视频内容,还可以存放或下载或其他虚拟现实内容,所述头戴式显示器设有主数据处理单元,在头戴式显示器能够正常工作的情况下,用于依据第一传感器组和第二传感器组的单组检测角度计算出用户头部运动角度的最终检测数据,依据用户头部运动角度对音频信号进行旋转并生成所述能够跟踪头部运动的虚拟环绕声。

[0028] 当所述头戴式显示器的电量不足时,可以直接传输音频到耳机中,即所述头戴式显示器不计算用户头部运动角度的最终检测结果,也不对音频信号进行旋转处理,而将其存储或从外部获得的音频直接传给耳机,并采用下列任意一种方式处理:

[0029] 方式一:由所述耳机的处理单元依据第一传感器组和第二传感器组的单组检测角度计算获得用户头部运动角度的最终检测数据,并依据所获得的用户头部运动角度进行音频信号的旋转,生成所述跟踪头部运动的虚拟环绕声;

[0030] 方式二:以第二传感器组的单组检测角度用户头部运动角度的最终检测数据,由所述耳机的处理单元进行音频信号的旋转,生成所述跟踪头部运动的虚拟环绕声,不关闭第一传感器组,第一传感器组的单独检测数据依然可以用于头戴式显示器视频等的控制;

[0031] 方式三:以第二传感器组的单组检测角度用户头部运动角度的最终检测数据,由所述耳机的处理单元进行音频信号的旋转,生成所述跟踪头部运动的虚拟环绕声,关闭第一传感器组或者关闭涉及第一传感器组的融合算法模块,通常可以选优关闭第一传感器组,在第一传感器硬件无法关闭时,可以关闭相关的融合算法模块,并将第二传感器组的单组检测数据传到所述头戴式显示器,以便用于视频等的控制。

[0032] 优选的,所述第一传感器组和第二传感器组在各自传感器融合计算中所用的时间常数可以不相同,例如,其中一个传感器组的时间常数为0.95,另一个传感器组的时间常数为0.99,在这种时间常数数值下,可以获得良好的听觉效果。

[0033] 在获得用户头部运动角度检测结果后,可以采用任意适宜的方式对音频信号进行旋转,以便消除消除因用户头部运动带来的方向失真,生成能够实时跟踪头部运动的虚拟环绕声,由耳机进行播放。

[0034] 下面给出申请人发明的一种涉及生成能够实时跟踪头部运动的虚拟环绕声的优选实施方法,主要为:依据所获得的用户头部运动角度的最终检测结果,生成用于消除用户头部旋转影响的旋转矩阵,根据第一音频信号(原音频),获取第一音频信号的调整参数,根据该调整参数,对第一音频信号进行调整得到第二音频信号(虚拟环绕声),根据第二音频信号和该旋转矩阵,生成跟踪头部运动的虚拟环绕声,从而可以提高虚拟环绕声的真实性,如果存储或从外部获得的音频已为第二音频信号或虚拟环绕声形式,则可以省略有第一音频信号向第二音频信号的转换。

[0035] 具体包括:

[0036] 步骤201:获取音频文件的第一音频信号和用户头部旋转的旋转角度;

[0037] 当用户通过耳机播放音频文件时,所用设备终端(头戴式显示器或耳机的相应处

理单元)获取音频文件的第一音频信号和用户头部旋转的旋转角度,该旋转角度采用通过前面所述的用户运动角度的最终检测数据。

[0038] 其中,第一音频信号可以为二阶B格式信号,B格式信号可以为三声道信号,也可以为四声道信号;如果B格式信号为三声道信号,则B格式信号包括W、X和Y;如果B格式信号为四声道信号,则B格式信号包括W、X、Y和Z。

[0039] W声道信号表示全方向声波,X声道信号、Y声道信号和Z声道信号表示沿三个互相垂直取向的声波;X声道信号表示听着从后至前水平布置,Y声道信号表示听着从右至左水平布置,Z声道信号表示听着向上垂直布置。

[0040] 步骤202:根据该旋转角度,生成旋转矩阵;

[0041] 旋转矩阵用于旋转虚拟环绕声,从而使得当用户头部旋转时,虚拟环绕声不根据用户头部的旋转而旋转,实现在现实生活中听音乐的效果。

[0042] 例如,虚拟环绕声的方向在前方,当用户头部向左旋转30度,则将该虚拟环绕声从用户头部旋转后的位置处向右旋转30度,从而实现虚拟环绕声的方向还是在原来的方向上。

[0043] 如果B格式信号为三声道信号,则旋转矩阵为
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix};$$

[0044] 如果B格式信号为四声道信号,则旋转矩阵为
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) & 0 \\ 0 & \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \theta \text{ 为该旋转角}$$

度。

[0045] 步骤203:根据第一音频信号,获取第一音频信号的调整参数;

[0046] 调整参数包括上混阶数和虚拟扬声器的拓扑结构,虚拟扬声器的拓扑结构包括虚拟扬声器的个数和每个虚拟扬声器的位置等。

[0047] 本步骤可以通过以下第一种方式或者第二种方式实现,对于第一种实现方式,本步骤可以通过以下步骤(1)和(2)实现,包括:

[0048] (1):根据第一音频信号,获取第一音频信号的录制场景;

[0049] 录制场景包括演唱会场景、商务会议场景或者自然环境场景等。

[0050] 本步骤可以通过以下步骤(1-1)和(1-2)实现,包括:

[0051] (1-1):分析第一音频信号,得到第一音频信号的内容;

[0052] 第一音频信号的内容至少包括指向性信号比重,还可以包括属性信息和/或主要声源的方向等;属性信息包括第一音频信号包括的对象、乐器类别和声音类别等。

[0053] 其中,分析第一音频信号,得到第一音频信号包括的指向性信号比重的步骤可以为:

[0054] 通过Direct-ambience signal decomposition(指向-环境信号分析)算法分析第一音频信号,得到第一音频信号包括的指向性信号的比重,也可以得到第一音频信号包括的非指向性信号的比重。

[0055] 例如,第一音频信号中只包含说话的声音,则第一音频信号听起来就会有有很强的方向性,则通过Direct-ambience signal decomposition算法分析第一音频信号,得到第一音频信号中的指向性信号的比重较大;再如,第一音频信号中包含噪音或者大量混响,则第一音频信号听起来方向性就不强,则通过Direct-ambience signal decomposition算法分析第一音频信号,得到第一音频信号中的指向性信号的比重较小。

[0056] 其中,分析第一音频信号,得到第一音频信号包括的指向性信号比重的步骤还可以通过以下步骤(A)至(C)实现,包括:

[0057] (A):对第一音频信号建立协方差矩阵;

[0058] 协方差矩阵 $\text{cov}(\mathbf{n}) = \sum_{\omega_i} \text{cov}(\omega_i, n)$,

[0059] $\text{cov}(\omega_i, n) = \alpha \text{cov}(\omega_i, n-1) + (1-\alpha) * S(\omega_i, n) * S^H(\omega_i, n)$ 。

[0060] 如果第一音频信号包括W、X和Y,

[0061] 则 $S(\omega_i, n) = [W(\omega_i, n) X(\omega_i, n) Y(\omega_i, n)]^T$,

[0062] 如果第一音频信号包括W、X、Y和Z,

[0063] 则 $S(\omega_i, n) = [W(\omega_i, n) X(\omega_i, n) Y(\omega_i, n) Z(\omega_i, n)]^T$ 。

[0064] 其中, ω_i 为第一音频信号的频率, n 为对时间轴上帧数的索引, $[\]^H$ 代表向量共轭转置; α 是平滑因子,且 α 可以事先设定或者根据第一音频信号的信号特性动态调整,例如, α 可以0.92。 ω_i 包括了所有感兴趣的频率并且 ω_i 可以根据需要进行设置并更改;例如, ω_i 为100-16000HZ。

[0065] 并且,可以设置每个 ω_i 的权重,则对第一音频信号建立协方差矩阵时,可以根据每个 ω_i 的权重计算第一音频信号的协方差矩阵,则协方差矩阵

[0066] $\text{cov}(n) = \sum \text{cov}(\omega_i, n) * \rho$, ρ 为 ω_i 的权重。

[0067] (B):对协方差矩阵进行特征分析,得到特征值;

[0068] 通过Matlab函数对协方差矩阵进行特征分析,得到 $[V, \Lambda] = \text{eigs}(\text{cov}(n))$ 。

[0069] 其中, V 是3*3的矩阵或者4*4的矩阵,该矩阵的每列代表 $\text{cov}(n)$ 的特征向量; Λ 包含了以降序排列的相应特征值。

[0070] (C):根据特征值,计算指向信号的比重。

[0071] 从特征值中选择最大特征值作为第一特征值,从除最大特征值之外的特征值中选择最大特征值作为第二特征值,根据第一特征值和第二特征值,通过以下公式(1)计算指向性信号的比重。

[0072]
$$\text{DRR} = 1 - \frac{\lambda_2}{\lambda_1}$$
 公式(1);

[0073] 其中,DRR为指向性信号的比重; λ_1 为第一特征值,且 λ_1 对应于直达声能量; λ_2 为第二特征值,且 λ_2 对应反射,回声,环境声等。DRR的值在[0,1]之间,DRR的值越小代表直达声比重越低,声场方向性越弱,也即指向性信号比重越低;DRR的值越大代表直达声比重越高,声场方向性越强,也即指向性信号比重越高。

[0074] 进一步地,通过direction of arrival(声源到达方向)算法分析第一音频信号,得到第一音频信号主要声源的方向。

[0075] 进一步地,通过Instrument classification(乐器分类)算法分析第一音频信号,

得到第一音频信号的乐器类别;通过Speechmusic classification(语音音乐分类)算法分析第一音频信号,得到第一音频信号的声音类别。

[0076] 进一步地,通过Object extraction(对象提取)算法提取第一音频信号包括的对象。

[0077] 例如,第一音频信号为一段语音,则通过Object extraction算法提取第一音频信号的对象为语音;再如,第一音频信号为一段雷声,则通过Object extraction算法提取第一音频信号的对象为雷声等;再如,第一音频信号为一段音乐,则通过Object extraction算法提取第一音频信号的对象为音乐等。

[0078] 进一步地,可以事先分析第一音频信号,得到第一音频信号的内容,将第一音频信号和第一音频信号的内容存储在音频信号和内容的对应关系中,则本步骤可以为:

[0079] 根据第一音频信号,从存储的音频信号和内容的对应关系中获取第一音频信号的内容。

[0080] 其中,存储的音频信号和内容的对应关系可以以元数据的形式存储在服务器中,并且可以直接将第一音频信号的内容嵌入第一音频信号中,也可以将第一音频信号的内容单独存放,建立内容文件夹,将第一音频信号的内容存储在该内容文件夹中,并建立第一音频信号和该内容文件夹的对应关系。

[0081] 可以在获取音频文件的第一音频时,获取第一音频信号的内容,也可以在本步骤中获取第一音频文件的内容。并且,从存储的音频信号和内容的对应关系中获取第一音频信号的内容,能够减轻运算负担,并且提高终端生成虚拟环绕声的效率。

[0082] (1-2):根据第一音频信号的内容,确定第一音频信号的录制场景。

[0083] 存储内容和录制场景的对应关系,相应的,本步骤可以为:

[0084] 根据第一音频信号的内容,从存储的内容和录制场景的对应关系中获取第一音频信号的录制场景。

[0085] 在本步骤中,也可以获取内容和录制场景的对应关系,存储内容和录制场景的对应关系;相应的,本步骤可以为:

[0086] 根据第一音频信号的内容,从存储的内容和录制场景的对应关系中获取第一音频信号的录制场景。

[0087] 其中,内容和录制场景的对应关系可以以元数据的形式存储,并且可以直接将第一音频信号的录制场景嵌入第一音频信号的内容中,也可以将第一音频信号的录制场景单独存放,建立录制场景文件夹,将第一音频信号的录制场景存储在该录制场景文件夹中,并建立第一音频信号的内容和该录制场景的对应关系。

[0088] 进一步地,将第一音频信号和第一音频信号的录制场景存储在音频信号和录制场景的对应关系中;从而再次播放第一音频信号时,不用通过以上方法确定第一音频信号的录制场景,直接从音频信号和录制场景的对应关系中获取第一音频信号的录制场景。

[0089] 例如,存储的指向性信号的比重大于0.5时,确定第一音频信号的录制场景为商务会议;指向性信号的比重小于0.5时,确定第一音频信号的录制场景为演唱会。

[0090] (2):根据录制场景,从录制场景和调整参数的对应关系中获取第一音频信号的调整参数。

[0091] 存储录制场景和调整参数的对应关系,则本步骤可以为:

[0092] 根据录制场景,从存储的录制场景和调整参数的对应关系中获取第一音频信号的调整参数。

[0093] 在本步骤中,也可以获取录制场景和调整参数的对应关系,存储录制场景和调整参数的对应关系;相应的,本步骤可以为:

[0094] 根据录制场景,从存储的录制场景和调整参数的对应关系中获取第一音频信号的调整参数。

[0095] 进一步地,将第一音频信号和第一音频信号的调整参数存储在音频信号和调整参数的对应关系中,从而再次播放第一音频信号时,不用通过以上方法确定第一音频信号的录制场景,在根据录制场景获取调整参数,而是直接从音频信号和调整参数的对应关系中获取第一音频信号的调整参数,从而缩短了获取第一音频信号的调整参数的获取时间,提高了获取效率。

[0096] 进一步地,对于第二种实现方式,本步骤可以为:

[0097] 存储音频信号和调整参数的对应关系,根据第一音频信号,从存储的音频信号和调整参数的对应关系中获取第一音频信号的调整参数。

[0098] 进一步地,音频信号和调整参数的对应关系可以以元数据的形式存储,并且可以直接将第一音频信号的调整参数嵌入第一音频信号中,也可以将第一音频信号的调整参数单独存放,建立调整参数文件夹,将第一音频信号的调整参数存储在调整参数文件中,并建立第一音频信号和该调整参数文件夹的对应关系。

[0099] 例如,第一音频信号中指向性信号的比重大于0.5,则上混阶数为3,且虚拟扬声器的拓扑结构包括6个虚拟扬声器;再如,第一音频信号中指向性信号的比重小于0.5,则上混阶数为1,且虚拟扬声器的拓扑结构包括4个虚拟扬声器。

[0100] 例如,第一音频信号的方向性声源分布在某一方向,比如声场内容为音乐会,声场集中在舞台方向也即正前方,则将虚拟扬声器的拓扑结构中的左前方和右前方扬声器的距离变大。

[0101] 虚拟扬声器的拓扑结构对虚拟环绕声的质量有重大的影响,并且对虚拟环绕声的影响根据第一音频信号的内容的不同而不同;例如,第一音频信号大部分都来自前方,则虚拟扬声器的拓扑结构会选择矩形结构,而不是方形结构。

[0102] 步骤204:根据该调整参数,对第一音频信号进行调整得到第二音频信号;

[0103] 调整参数包括上混阶数和虚拟扬声器的拓扑结构,则本步骤可以通过以下步骤(1)和(2)实现,包括:

[0104] (1):根据上混阶数,将第一音频信号进行上混处理得到第四音频信号;

[0105] 其中,本步骤为现有技术,在此不再详细说明。

[0106] (2):根据虚拟扬声器的拓扑结构,将第四音频信号进行环绕处理得到第二音频信号。

[0107] 将第四音频信号依次经过虚拟扬声器的拓扑结构中包括的虚拟扬声器,从而实现将第四音频信号进行环绕处理,得到第二音频信号。

[0108] 例如,第一音频信号为 $\begin{bmatrix} W_1 \\ X_1 \\ Y_1 \end{bmatrix}$,则调整后的第二音频信号为 $\begin{bmatrix} W_2 \\ X_2 \\ Y_2 \end{bmatrix}$;再如,第一音频

信号为 $\begin{bmatrix} W_1 \\ X_1 \\ Y_1 \\ Z_1 \end{bmatrix}$ ，则调整后的第二音频信号为 $\begin{bmatrix} W_2 \\ X_2 \\ Y_2 \\ Z_2 \end{bmatrix}$ 。

[0109] 步骤205:根据旋转矩阵,将第二音频信号进行旋转得到第三音频信号;

[0110] 将旋转矩阵和第二音频信号进行乘法运算,得到第三音频信号。

[0111] 例如,第二音频信号为 $\begin{bmatrix} W_2 \\ X_2 \\ Y_2 \end{bmatrix}$, 旋转矩阵为 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) - \sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}$, 则第三音频信号为

$\begin{bmatrix} W_3 \\ X_3 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) - \sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} W_2 \\ X_2 \\ Y_2 \end{bmatrix}$; 再如,第二音频信号为 $\begin{bmatrix} W_2 \\ X_2 \\ Y_2 \\ Z_2 \end{bmatrix}$, 旋转矩阵为

$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) - \sin(\theta) & 0 & 0 \\ 0 & \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, 则第三音频信号为 $\begin{bmatrix} W_3 \\ X_3 \\ Y_3 \\ Z_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) - \sin(\theta) & 0 & 0 \\ 0 & \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} W_2 \\ X_2 \\ Y_2 \\ Z_2 \end{bmatrix}$ 。

[0112] 步骤206:根据第三音频信号,将第三音频信号转换为虚拟扬声器阵列信号;

[0113] 获取虚拟扬声器矩阵,将虚拟扬声器矩阵和第三音频信号进行矩阵乘法,得到虚拟扬声器阵列信号。

[0114] 例如,虚拟扬声器阵列为 $\begin{bmatrix} G_{w1} & G_{x1} & G_{y1} \\ G_{w2} & G_{x2} & G_{y2} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ G_{wN} & G_{xN} & G_{yN} \end{bmatrix}$, 则虚拟扬声器阵列信号为 $\begin{bmatrix} L_1 \\ L_2 \\ \cdot \\ L_N \end{bmatrix}$

$= \begin{bmatrix} G_{w1} & G_{x1} & G_{y1} \\ G_{w2} & G_{x2} & G_{y2} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ G_{wN} & G_{xN} & G_{yN} \end{bmatrix} \begin{bmatrix} W_3 \\ X_3 \\ Y_3 \end{bmatrix}$; 再如,虚拟扬声器阵列为 $\begin{bmatrix} G_{w1} & G_{x1} & G_{y1} & G_{z1} \\ G_{w2} & G_{x2} & G_{y2} & G_{z2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ G_{wN} & G_{xN} & G_{yN} & G_{zN} \end{bmatrix}$, 则虚拟扬声器阵

$$\text{列信号为} \begin{bmatrix} L_1 \\ L_2 \\ \dots \\ L_N \end{bmatrix} = \begin{bmatrix} G_{w1} & G_{x1} & G_{y1} & G_{z1} \\ G_{w2} & G_{x2} & G_{y2} & G_{z2} \\ \cdot & \cdot & \cdot & \cdot \\ G_{wN} & G_{xN} & G_{yN} & G_{zN} \end{bmatrix} \begin{bmatrix} W_3 \\ X_3 \\ Y_3 \\ Z_3 \end{bmatrix} \circ$$

[0115] 其中,N为虚拟扬声器拓扑结构中包括的虚拟扬声器的数目。

[0116] 步骤207:将虚拟扬声器阵列信号通过头相关变换函数滤波器进行滤波,得到虚拟环绕声。

[0117] 头相关变换函数滤波器用于将虚拟扬声器阵列信号转换为二路立体声,也即双耳信号,则本步骤可以为:

[0118] 获取头相关函数变换函数滤波器对应的二路立体声矩阵,将该二路立体矩阵和虚拟扬声器阵列信号进行矩阵乘法,得到虚拟环绕声。

[0119] 例如,二路立体声矩阵为 $\begin{bmatrix} H_{1L} & H_{2L} & \dots & H_{NL} \\ H_{1R} & H_{2R} & \dots & H_{NR} \end{bmatrix}$, 则虚拟环绕声为

$$\begin{bmatrix} L \\ R \end{bmatrix} = \begin{bmatrix} H_{1L} & H_{2L} & \dots & H_{NL} \\ H_{1R} & H_{2R} & \dots & H_{NR} \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \dots \\ L_N \end{bmatrix} = \begin{bmatrix} F_{WL} & F_{XL} & F_{YL} \\ F_{WR} & F_{XR} & F_{YR} \end{bmatrix} \begin{bmatrix} W_1 \\ X_1 \\ Y_1 \end{bmatrix}; \text{或者虚拟环绕声为}$$

$$\begin{bmatrix} L \\ R \end{bmatrix} = \begin{bmatrix} H_{1L} & H_{2L} & \dots & H_{NL} \\ H_{1R} & H_{2R} & \dots & H_{NR} \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \dots \\ L_N \end{bmatrix} = \begin{bmatrix} F_{WL} & F_{XL} & F_{YL} & F_{ZL} \\ F_{WR} & F_{XR} & F_{YR} & F_{ZR} \end{bmatrix} \begin{bmatrix} W_1 \\ X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \circ$$

[0120] 本发明公开的各优选和可选的技术手段,除特别说明外及一个优选或可选技术手段为另一技术手段的进一步限定外,均可以任意组合,形成若干不同的技术方案。