



(12) 发明专利申请

(10) 申请公布号 CN 105578378 A

(43) 申请公布日 2016. 05. 11

(21) 申请号 201511023469. 8

(22) 申请日 2015. 12. 30

(71) 申请人 深圳市有信网络技术有限公司

地址 518000 广东省深圳市南山区大冲商务中心 D 栋 10 层 1002 室

(72) 发明人 蒋明

(74) 专利代理机构 广州市越秀区哲力专利商标

事务所 (普通合伙) 44288

代理人 李悦 张鹏

(51) Int. Cl.

H04S 1/00(2006. 01)

H04S 7/00(2006. 01)

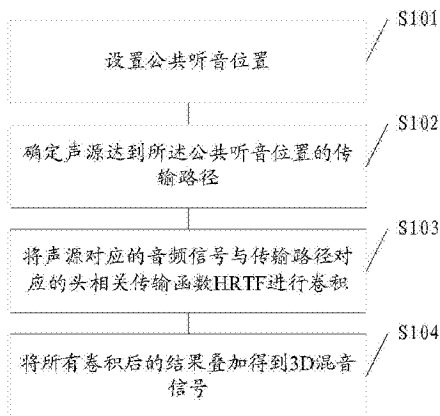
权利要求书1页 说明书5页 附图2页

(54) 发明名称

一种 3D 混音方法及装置

(57) 摘要

本发明涉及 3D 混音技术领域, 提供了一种 3D 混音方法, 方法包括以下步骤: 设置公共听音位置, 确定声源到达所述公共听音位置的传输路径, 将声源对应的音频信号与传输路径对应的头相关传输函数 HRTF 进行卷积, 将所有卷积后的结果叠加得到 3D 混音信号, 本发明的设置公共听音位置, 将声源对应的音频信号与传输路径对应的头相关传输函数 HRTF 进行卷积时, 计算量大为减少。



1. 一种3D混音方法,其特征在于,所述方法包括以下步骤:
设置公共听音位置;
确定声源到达所述公共听音位置的传输路径;
将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积;
将所有卷积后的结果叠加得到3D混音信号。
2. 如权利要求1所述的3D混音方法,其特征在于,所述确定所有发声位置达到所述公共听音位置的传输路径,包括:
根据声源到达公共听音位置的直达路径和反射路径确定所有传输路径。
3. 如权利要求2所述的3D混音方法,其特征在于,将所有卷积后的结果叠加得到3D混音信号,之前还包括:
将卷积后的结果做衰减和延时处理。
4. 如权利要求3所述的3D混音方法,其特征在于,将所有卷积后的结果叠加得到3D混音信号,具体为:
将所有做衰减和延时处理后的卷积结果叠加得到3D混音信号。
5. 如权利要求2所述的3D混音方法,其特征在于,所述将所有卷积后的结果做衰减和延时处理,具体为将卷积后的结果乘以衰减系数,并根据延迟时间对卷积后的结果进行延时处理,所述衰减系数 $\text{decay} = 0.5 * \log_2(L/L_d)$,所述延迟时间 $\text{delay} = (L_d - L)/v$,其中所述L表示反射路径的路径总长,所述 L_d 表示直达路径的路径总长,所述v表示声音速度。
6. 一种3D混音装置,其特征在于,所述装置包括:
设置模块,用于设置公共听音位置;
确定模块,用于确定声源到达所述公共听音位置的传输路径;
卷积模块,用于将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积;
叠加模块,用于将所有卷积后的结果叠加得到3D混音信号。
7. 如权利要求6所述的3D混音装置,其特征在于,所述确定模块,还具体用于:
根据声源到达公共听音位置的直达路径和反射路径确定所有传输路径。
8. 如权利要求7所述的3D混音装置,其特征在于,装置还包括:
衰减和延迟模块,用于将卷积后的结果做衰减和延时处理。
9. 如权利要求8所述的3D混音装置,其特征在于,所述卷积模块,还用于:
将所有做衰减和延时处理后的卷积结果叠加得到3D混音信号。
10. 如权利要求7所述的3D混音装置,其特征在于,所述衰减和延时模块,还用于将卷积后的结果乘以衰减系数,并根据延迟时间对卷积后的结果进行延时处理;
所述衰减系数 $\text{decay} = 0.5 * \log_2(L/L_d)$,所述延迟时间 $\text{delay} = (L_d - L)/v$,其中所述L表示反射路径的路径总长,所述 L_d 表示直达路径的路径总长,所述v表示声音速度。

一种3D混音方法及装置

技术领域

[0001] 本发明涉及3D混音技术领域,尤其涉及一种3D混音方法及装置。

背景技术

[0002] 在多人参与的电话会议中,每个人经电话麦克风采集到的信号都是单声道的。简单的将所有人的声音混合,会有一些问题:首先,在多个人同时讲话的场景下,互相叠加的声音会对听音者造成听觉和理解上的困难;其次,没有方位信息的混音会缺乏临场感。但是如果引入3D混音,这些问题就会得到解决。HRTF卷积是一种适应性很好的3D混音方法,将单声道音频信号与指定的HRTF函数卷积,就可以获得能够用耳机重放的立体声,混音让听音者通过耳机重放,听音者才能感觉到每一个说话人都在自己身边的不同位置上。在现有技术中,对于处理有N个人参加会议的3D混音方法,对于每一个人来说,其他N-1个人都是说话人,而自己是唯一的听音者,为了让每一个人听到3D混音后的效果,必须指定N-1对HRTF并且用它们与其他每一个人的声音信号卷积。由于对每一个人来说,其他所有人都分布在不同的位置,因此,为了让所有人听到3D效果,就要有 $N*(N-1)$ 对HRTF卷积。也就是现有技术中HRTF卷积方法的计算量需要N的平方的数量级,计算复杂。

发明内容

[0003] 本发明的目的在于提供一种3D混音方法,旨在解决现有的3D混音中HRTF卷积时,计算量复杂的问题。

[0004] 本发明提供的一种3D混音方法,方法包括以下步骤:

[0005] 设置公共听音位置;

[0006] 确定声源到达公共听音位置的传输路径;

[0007] 将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积;

[0008] 将所有卷积后的结果叠加得到3D混音信号。

[0009] 进一步的,确定所有发声位置达到所述公共听音位置的传输路径,包括:

[0010] 根据声源到达公共听音位置的直达路径和反射路径确定所有传输路径。

[0011] 进一步的,将所有卷积后的结果叠加得到3D混音信号,之前还包括:

[0012] 将卷积后的结果做衰减和延时处理。

[0013] 进一步的,将所有卷积后的结果叠加得到3D混音信号,具体为:

[0014] 将所有做衰减和延时处理后的卷积结果叠加得到3D混音信号。

[0015] 进一步的,将所有卷积后的结果做衰减和延时处理,具体为将卷积后的结果乘以衰减系数,并根据延迟时间对卷积后的结果进行延时处理,所述衰减系数 $\text{decay} = 0.5 * \log_2(L/L_d)$,所述延迟时间 $\text{delay} = (L_d - L) / v$,其中所述L表示反射路径的路径总长,所述 L_d 表示直达路径的路径总长,所述v表示声音速度。

[0016] 本发明还提供了一种3D混音装置,该装置包括:

[0017] 设置模块,用于设置公共听音位置;

- [0018] 确定模块,用于确定声源到达所述公共听音位置的传输路径;
- [0019] 卷积模块,用于将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积;
- [0020] 叠加模块,用于将所有卷积后的结果叠加得到3D混音信号。
- [0021] 进一步的,确定模块,还具体用于:
- [0022] 根据声源到达公共听音位置的直达路径和反射路径确定所有传输路径。
- [0023] 进一步的,3D混音装置还包括:
- [0024] 衰减和延迟模块,用于将卷积后的结果做衰减和延时处理。
- [0025] 进一步的,卷积模块,还用于:
- [0026] 将所有做衰减和延时处理后的卷积结果叠加得到3D混音信号。
- [0027] 进一步的,衰减和延时模块,还用于将卷积后的结果乘以衰减系数,并根据延迟时间对卷积后的结果进行延时处理;
- [0028] 所述衰减系数 $decay = 0.5 * \log_2(L/L_d)$,所述延迟时间 $delay = (L_d - L)/v$,其中所述L表示反射路径的路径总长,所述 L_d 表示直达路径的路径总长,所述v表示声音速度。
- [0029] 本发明的3D混音方法及装置有益效果:本发明的3D混音方法设置公共听音位置,让3D混音中将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积时,HRTF卷积方法的计算量为N的数量级,大大减少了运算量。

附图说明

- [0030] 图1是本发明一实施例提供的一种3D混音方法的流程示意图;
- [0031] 图2是本发明一实施例提供的声源A到达公共听音位置T处的几条主要传输路径示意图;
- [0032] 图3是本发明又一实施例提供的3D混音方法的流程示意图;
- [0033] 图4是本发明一实施例提供的一种3D混音装置。

具体实施方式

- [0034] 为了使本发明的目的、技术方案及优点更加清楚明白,下面结合附图和具体实施方式对本发明作进一步详细说明。
- [0035] 图1是本发明一实施例提供的一种3D混音方法的流程示意图,参考图1,本发明实施例的3D混音方法,包括以下步骤:
- [0036] 步骤S101,设置公共听音位置。
- [0037] 在3D混音技术中,由于是虚拟混音,每一个人说话和听音在不同的位置进行,这样的设定是可行的。例如,对于参会者A,他在 A_s 处说话而在 A_l 处听音;对于参会者B,他在 B_s 处说话在 B_l 处听音,等等。现在取所有人的听音位置相同,假设在T处,也就是 A_l , B_l 以及其他人的听音位置都是同一个地方T处,那么不管是参会者A,还是参会者B,在听音的时候,说话人相对于T处位置都是相同的。
- [0038] 步骤S102,确定声源到达公共听音位置的传输路径。
- [0039] 包括:根据声源到达公共听音位置的直达路径和反射路径确定所有传输路径。
- [0040] 图2是本发明一实施例提供的声源A处到达公共听音位置T处的几条主要传输路径

示意图。参考图2,参会者A实际上说话时,声音达到公共听音位置T处的传输路径有无数条,先确定几条主要的承载绝大部分声音能量的传输路径,比如在一个长方体的房间里,长方体有六个面:U1U2U3U4、U5U6U7U8、U1U2U7U8、U3U4U5U6、U1U4U6U7、U2U3U5U8。假设参会者A(也即说话者)位置为U1U2U3U4对应的中心为位置,公共听音位置T设置在面U5U6U7U8的中心,K、M、Q、P分别为面U1U2U7U8、U3U4U5U6、U1U4U6U7、U2U3U5U8的中心。则参会者A说话的声音从A处传输到公共听音位置T处,有5条主要的路径,声源A处到达公共听音位置T处的直达路径AT,声源到达公共听音位置的反射路径有:从左侧面反射的路径AKT、从右侧面反射的路径AMT、从下面反射的路径APT、从上面反射的路径AQT。依此方法,还可以测出其他的一些路径,比如从A处经过左侧面反射到右侧面、再从右侧面反射到达T处的路径,或声音从A处经过下面反射到上面、再从上面反射到达T处的路径,或声音从A处经过左侧面反射到右侧面、再从右侧面反射到左侧面、再从左侧面反射到达T处,等等,其他的路径就不一一列举。但是,由于直达路径AT、反射的路径AKT、AMT、APT、AQT这5条传声路径承载了绝大部分的声音能量,因此,考虑房间混音是只需要考虑他们就可以了。

[0041] 刚才列举的例子,是一个有规则的长方体的环境中,在其他的环境中也同样可以确定声源到达公共听音位置的传输路径。包括:根据声源到达公共听音位置的直达路径和反射路径确定所有传输路径。

[0042] 依上面的实施例也可以确定其他参会者(说话者)到公共听音位置的传输路径,包括:根据声源到达公共听音位置的直达路径和反射路径确定所有传输路径,也不限于只是在长方体的环境中的所有从声源到达公共听音位置的传输路径。

[0043] 步骤S103,将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积。

[0044] 从步骤S102中确定了每个声源到达公共听音位置的传输路径,也就能找到相应的传输路径对应的头相关传输函数HRTF,本发明实施例的3D混音方法将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积,从而得到相应的3D虚拟混音信号。

[0045] 本发明实施例中,HRTF函数是通过查找HRIR表,HRIR表示根据声源位置相对于听音者的正前方方向的角度来查找的。HRTF表征的是声源相对于听音者正前方方向的角度。这个角度表述为一个2维向量 $\langle x, y \rangle$ 其中x指水平角度,y指俯仰角度,取值范围都是0-360度。例如 $\langle 30, 60 \rangle$ 指的是声源位置在听音者正前方偏左30度,偏上60。注意偏左90度是正左方,180是后方,270是正右。同理偏上90度是正上,270度是正下。而HRIR表的获取方法可以通过实验用人工头耳内录音获取。例如在距离人工头单位长度的远近,角度为 $\langle x, y \rangle$ 的地方发出一个脉冲声音,记录下左右耳的响应,得到HRIR表的第m行k列的HRTF。依次测完所有需要的 $\langle x, y \rangle$,做成了一个表,就是HRIR表。例如,以一度为单位来测量,就得到360行360列的表,那么对应角度 $\langle x, y \rangle$ 的HRTF就在表的x行y列(当然实际上不需要这么多)。查表时,就只需要知道角度 $\langle x, y \rangle$,直接找到对应表中相应位置的HRTF就可以了。实际中,如果表中没有 $\langle x, y \rangle$ 的HRTF,则需要用插值的方法用靠近这个角度的其他角度的数据估算出一个该角度的HRTF。

[0046] 步骤S104,将所有卷积后的结果叠加得到3D混音信号。

[0047] 每加入一个参会者,把所有卷积后的结果叠加得到3D混音信号。例如,设置公共听音位置后,确定有三个声源:参会者A、B、C,参会者A发出的声音有五条传输路径到达公共听

音位置,参会者B发出的声音有四条传输路径到达公共听音位置,参会者C发出的声音有六条传输路径到达公共听音位置,将声源A对应的音频信号与传输的五条路径对应的头相关传输函数HRTF进行五次卷积,将声源B对应的音频信号与传输的四条路径对应的头相关传输函数HRTF进行四次卷积,将声源C对应的音频信号与传输的六条路径对应的头相关传输函数HRTF进行六次卷积,将一起十五次卷积后的结果叠加得到最终的3D混音信号。

[0048] 本发明实施例中,由于设置公共听音位置,N人参与的会议,只需要N对HRTF卷积。这对于运算量的节省具有非常大的意义。可以看出,传统方法的计算量需要N的平方的数量级,而本发明实施例的方法仅需要N的数量级。

[0049] 图3是本发明又一实施例提供的3D混音方法的流程示意图,参考图3,该方法包括以下步骤:

[0050] 步骤S301,设置公共听音位置。

[0051] 步骤S302,确定声源到达公共听音位置的传输路径。

[0052] 步骤S303,将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积。

[0053] 上述步骤S301-步骤S303与前述实施例的步骤S101-步骤S103相同,在此不再赘述。

[0054] 步骤S304,将卷积后的结果做衰减和延时处理。

[0055] 本发明实施例根据实际现象考虑的,实际场景中,两个人对话,说话人用同样大小的声音说出来,听音者听到的声音大小与两个人的距离相关,距离越大听到的声音越小,距离越大,听到声音的时间离声音发出的时间越晚。从步骤S102中确定了每个声源到达公共听音位置的传输路径,单个声源传输路径是有无数条,每条路径到达听音者的长度是不一样的,到达听音者的音量大小也不样,到达听音者的时间也不一样,这就存在一个声音衰减和延时的问题,本发明实施例将所有卷积后的结果做衰减和延时处理,比如参会者A处发声经五条路径到达公共听音位置T处,直达路径AT的长度就是AT的直线距离,其他反射路径AKT的长度为线段AK加线段KT、其他反射路径AMT的长度为线段AM加线段MT、反射路径APT的长度为线段AP加线段PT、反射路径AQT的长度为线段AQ加线段QT,直达路径AT的长度最短,声音经路径AT传输到公共位置时,声音最大,而其他反射路径传输到公共位置T处时,声音相对小一些,本发明实施例根据直达路径AT和反射路径AKT、AMT、APT、AQT传输到公共位置时的声音大小不同,将所有卷积后的结果做衰减和延时处理,具体为将卷积后的结果乘以衰减系数,并根据延迟时间对卷积后的结果进行延时处理,衰减系数 $\text{decay} = 0.5 * \log_2(L/L_d)$,延迟时间 $\text{delay} = (L_d - L)/v$,其中L表示反射路径的路径总长,L_d表示直达路径的路径总长,所述v表示声音速度。延时处理是指对同一声源,由于有直达路径与有多条反射路径,其到达公共听音位置的时间是不一样的,在做叠加时,需要将对应的路径进行延迟,统一到一个时间点,再做叠加。

[0056] 步骤S305:将所有做衰减和延时处理后的卷积结果叠加得到3D混音信号。

[0057] 具体叠加方法与前述实施例步骤S104相同,在此不再赘述。

[0058] 本发明实施例中,通过衰减和延迟处理,增强了HRTF卷积效果,也增强了3D混音的效果。

[0059] 图4是本发明一实施例提供的一种3D混音装置,该装置与前述一实施例提供的方

法相对应。

[0060] 参考图4,本发明一实施例提供了一种3D混音装置包括:

[0061] 设置模块1,用于设置公共听音位置;

[0062] 确定模块2,用于确定声源到达公共听音位置的传输路径;

[0063] 卷积模块3,用于将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积;

[0064] 叠加模块4,用于将所有卷积后的结果叠加得到3D混音信号。

[0065] 进一步的,确定模块2,还具体用于:

[0066] 根据声源到达公共听音位置的直达路径和反射路径确定所有传输路径。

[0067] 进一步的,装置还包括:

[0068] 衰减和延迟模块,用于将卷积后的结果做衰减和延时处理。

[0069] 进一步的,卷积模块3,还用于:

[0070] 将所有做衰减和延时处理后的卷积结果叠加得到3D混音信号。

[0071] 进一步的,衰减和延时模块,还用于将卷积后的结果乘以衰减系数,并根据延迟时间对卷积后的结果进行延时处理;

[0072] 所述衰减系数 $\text{decay} = 0.5 * \log_2(L/L_d)$,所述延迟时间 $\text{delay} = (L_d - L)/v$,其中所述L表示反射路径的路径总长,所述 L_d 表示直达路径的路径总长,所述v表示声音速度。

[0073] 本实施例中的装置与前述实施例中的3D混音方法是基于同一发明构思下的两个方面,在前面已经对方法实施过程作了详细的描述,所以本领域技术人员可根据前述描述清楚地了解本实施中的系统的结构及实施过程,为了说明书的简洁,在此就不再赘述。

[0074] 本发明实施例提供的一个或多个技术方案,至少有以下技术效果:

[0075] 1)由于设置了公共听音位置,一个N人参与的会议,将声源对应的音频信号与传输路径对应的头相关传输函数HRTF进行卷积的次数只用N次,比现有技术中用 $N*(N-1)$ 次的运算,运算量可以大大的减小。

[0076] 2)由于将卷积后的结果做衰减和延时处理,增强了HRTF卷积效果,也增强了3D混音的效果。

[0077] 对本领域的技术人员来说,可根据以上描述的技术方案以及构思,做出其它各种相应的改变以及形变,而所有的这些改变以及形变都应该属于本发明权利要求的保护范围之内。

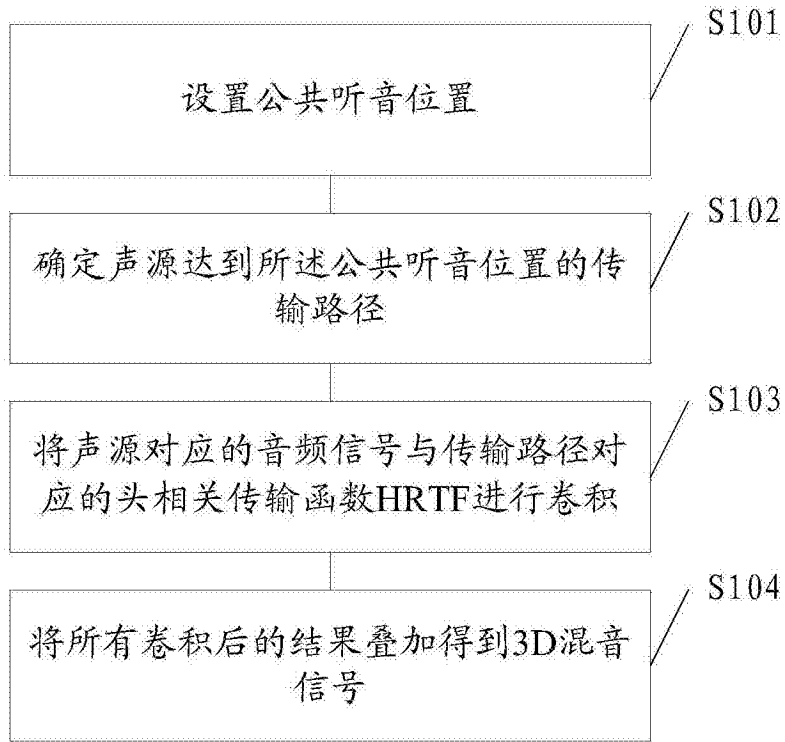


图1

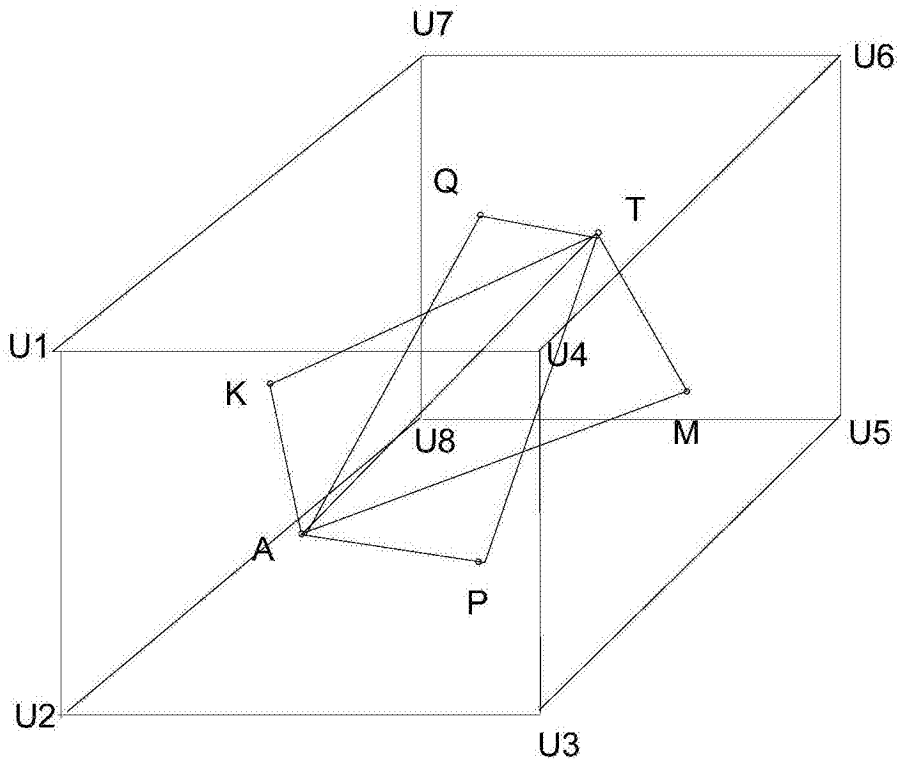


图2



图3

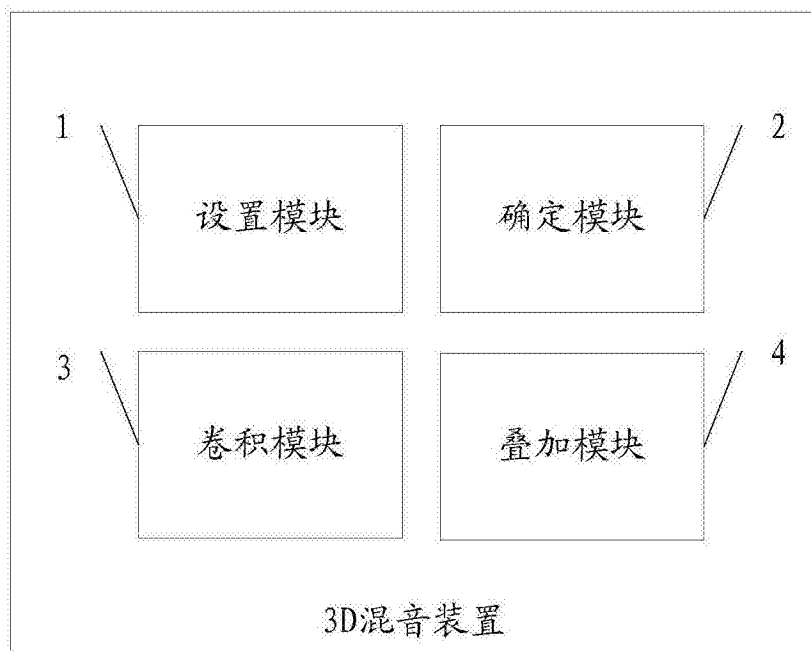


图4