一种音频编解码方法和装置

本申请要求于2020年11月30日提交中国专利局、申请号为202011377320.0、发明名称为“一种音频编解码方法和装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及音频编解码技术领域，尤其涉及一种音频编解码方法和装置。

背景技术

三维音频技术是对真实世界中的声音事件和三维声场信息进行获取、处理、传输和渲染回放的音频技术。三维音频技术使声音具有强烈的空间感、包围感及沉浸感，给人以“声临其境”的非凡听觉体验。高阶立体混响(higher order ambisonics，HOA) 技术具有在录制、编码与回放阶段与扬声器布局无关的性质和 HOA 格式数据的可旋转回放特性，在进行三维音频回放时具有更高的灵活性，因而也得到了更为广泛的关注和研究。

为了实现更好的音频听觉效果，HOA 技术需要大量的数据量用于记录更详细的声音场景的信息。虽然这种基于场景的三维音频信号采样和存储更加利于音频信号空间信息的保存和传输，但随着 HOA 阶数的增加将会产生更多的数据，大量的数据造成传输和存储的困难，因此需要对 HOA 信号进行编解码。

目前存在一种多声道数据的编解码方法，包括：在编码端，通过核心编码器（例如 16 声道的编码器）直接对原始场景音频信号的每一个声道进行编码，然后输出码流。在解码端，通过核心解码器（例如 16 声道的解码器）对码流进行解码，以得到解码场景音频信号的每一个声道。

上述多声道编解码方法，需要根据原始场景音频信号的声道数适配相应的编解码器，且随着声道数增加，压缩码流存在数据量大、带宽占用高的问题。

发明内容

本申请实施例提供了一种音频编解码方法和装置，用于减少编解码的数据量，以提高编解码效率。

为解决上述技术问题，本申请实施例提供以下技术方案：

第一方面，本申请实施例提供一种音频编码方法，包括：

根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器；

根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号；

对所述第一虚拟扬声器信号进行编码，以得到码流。

在本申请实施例中，根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器；根据当前场景音频信号和第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号；对第一虚拟扬声器信号进行编码，以得到码流。由于本申请实施例中可以根据

第一场景音频信号和第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，音频编码端对该第一虚拟扬声器信号进行编码，而不再直接对第一场景音频信号进行编码，本申请实施例中根据第一场景音频信号选择出第一目标虚拟扬声器，基于该第一目标虚拟扬声器生成的第一虚拟扬声器信号可以表示空间中听音人所在的位置声场，该位置声场尽可能的接近录制第一场景音频信号时的原始声场，保证了音频编码端的编码质量，且对第一虚拟扬声器信号和残差信号进行编码以得到码流，该第一虚拟扬声器信号的编码数据量与第一目标虚拟扬声器有关，而与第一场景音频信号的声道个数无关，减少了编码数据量，提高编码效率。

在一种可能的实现方式中，所述方法还包括：

根据所述虚拟扬声器集合从所述当前场景音频信号中获取主要声场成分；

所述根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器包括：

根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器。

在上述方案中，虚拟扬声器集合中的每个虚拟扬声器对应一个声场成分，则根据主要声场成分从虚拟扬声器集合中选择出第一目标虚拟扬声器，例如主要声场成分对应的虚拟扬声器就是编码端选择出的第一目标虚拟扬声器。本申请实施例中，编码端通过主要声场成分可以选择出第一目标虚拟扬声器，解决了编码端需要确定第一目标虚拟扬声器的问题。

在一种可能的实现方式中，所述根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器，包括：

根据所述主要声场成分从高阶立体混响 HOA 系数集合中选择出与所述主要声场成分对应的 HOA 系数，所述 HOA 系数集合中的 HOA 系数与所述虚拟扬声器集合中的虚拟扬声器一一对应；

确定所述虚拟扬声器集合中与所述主要声场成分对应的 HOA 系数对应的虚拟扬声器为所述第一目标虚拟扬声器。

在上述方案中，编码端中根据虚拟扬声器集合预先配置 HOA 系数集合，HOA 系数集合中的 HOA 系数与虚拟扬声器集合中的虚拟扬声器之间的一一对应关系，因此根据主要声场成分选择出 HOA 系数之后，再根据上述一一对应关系从虚拟扬声器集合中查找与主要声场成分对应的 HOA 系数对应的目标虚拟扬声器，该查找出的目标虚拟扬声器即为第一目标虚拟扬声器，解决了编码端需要确定第一目标虚拟扬声器的问题。

在一种可能的实现方式中，所述根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器，包括：

根据所述主要声场成分获取所述第一目标虚拟扬声器的配置参数；

根据所述第一目标虚拟扬声器的配置参数生成所述第一目标虚拟扬声器对应的 HOA 系数；

确定所述虚拟扬声器集合中所述第一目标虚拟扬声器对应的 HOA 系数对应的虚拟扬声器为所述目标虚拟扬声器。

在上述方案中，编码端在获取到主要声场成分之后，可以根据该主要声场成分可用于确定第一目标虚拟扬声器的配置参数，例如主要声场成分是多个声场成分中取值最大的一

个或几个声场成分，或主要声场成分可以是多个声场成分中方向占优的一个或几个声场成分，该主要声场成分可用于确定出当前场景音频信号匹配的第一目标虚拟扬声器，第一目标虚拟扬声器配置有相应的属性信息，使用第一目标虚拟扬声器的配置参数都可以生成该第一目标虚拟扬声器的 HOA 系数，HOA 系数的生成过程可以通过 HOA 算法来实现，此处不再详细说明。虚拟扬声器集合中每个虚拟扬声器都对应有 HOA 系数，因此可以根据每个虚拟扬声器对应的 HOA 系数从虚拟扬声器集合中选择出第一目标虚拟扬声器，解决了编码端需要确定第一目标虚拟扬声器的问题。

在一种可能的实现方式中，所述根据所述主要声场成分获取所述第一目标虚拟扬声器的配置参数，包括：

根据音频编码器的配置信息确定所述虚拟扬声器集合中的多个虚拟扬声器的配置参数；

根据所述主要声场成分从所述多个虚拟扬声器的配置参数中选择出所述第一目标虚拟扬声器的配置参数。

在上述方案中，音频编码器中可以预先存储多个虚拟扬声器各自的配置参数，每个虚拟扬声器的配置参数可以是通过音频编码器的配置信息确定，音频编码器是指前述的编码端，该音频编码器的配置信息，包括且不限于：HOA 阶数、编码比特率等。音频编码器的配置信息可以用于确定虚拟扬声器的个数，以及每个虚拟扬声器的位置参数，解决了编码端需要确定虚拟扬声器的配置参数的问题。举例说明如下，若编码比特率较低时可以配置较少数量的虚拟扬声器，若编码比特率较高时可以配置多个数量的虚拟扬声器。又如虚拟扬声器的 HOA 阶数可以等于音频编码器的 HOA 阶数。不限定的是，本申请实施例中，除了通过音频编码器的配置信息确定多个虚拟扬声器各自的配置参数之外，还可以根据用户自定义信息多个虚拟扬声器各自的配置参数，例如，用户可以自定义虚拟扬声器的位置、HOA 阶数、虚拟扬声器的个数等。

在一种可能的实现方式中，所述第一目标虚拟扬声器的配置参数包括：所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息；

所述根据所述第一目标虚拟扬声器的配置参数生成所述第一目标虚拟扬声器对应的 HOA 系数，包括：

根据所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息确定所述第一目标虚拟扬声器对应的 HOA 系数。

在上述方案中，使用每个虚拟扬声器的位置信息和 HOA 阶数信息都可以生成该虚拟扬声器的 HOA 系数，HOA 系数的生成过程可以通过 HOA 算法来实现，解决了编码端需要确定第一目标虚拟扬声器的 HOA 系数的问题。

在一种可能的实现方式中，所述方法还包括：

对所述第一目标虚拟扬声器的属性信息进行编码，并写入所述码流。

在上述方案中，编码端除了对虚拟扬声器进行编码，还可以对第一目标虚拟扬声器的属性信息进行编码，并将编码后的第一目标虚拟扬声器的属性信息写入到码流中，此时得到的码流中可以包括：编码后的虚拟扬声器和编码后的第一目标虚拟扬声器的属性信息。本申请实施例中码流中可以携带编码后的第一目标虚拟扬声器的属性信息，使得解码端通过解码码流，就可以确定出第一目标虚拟扬声器的属性信息，便于解码端的音频解码。

在一种可能的实现方式中，所述当前场景音频信号包括：待编码高阶立体混响 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的 HOA 系数；

所述根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，包括：

对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

在上述方案中，以当前场景音频信号为待编码 HOA 信号为例，编码端首先确定第一目标虚拟扬声器的 HOA 系数，例如编码端根据主要声场成分从 HOA 系数集合中选择出 HOA 系数，该选择出的 HOA 系数就是第一目标虚拟扬声器的 HOA 系数，编码端获取到待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数之后，根据待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数可以生成第一虚拟扬声器信号，其中，待编码 HOA 信号可以采用第一目标虚拟扬声器的 HOA 系数进行线性组合得到，第一虚拟扬声器信号的求解可以被转换为对线性组合的求解问题。

在一种可能的实现方式中，所述当前场景音频信号包括：待编码高阶立体混响 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的位置信息；

所述根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，包括：

根据所述第一目标虚拟扬声器的位置信息获取所述第一目标虚拟扬声器对应的 HOA 系数；

对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

在上述方案中，第一目标虚拟扬声器的属性信息可以包括：第一目标虚拟扬声器的位置信息，编码端预先存储虚拟扬声器集合中每个虚拟扬声器的 HOA 系数，编码端还存储有每个虚拟扬声器的位置信息，虚拟扬声器的位置信息和该虚拟扬声器的 HOA 系数之间存在对应关系，因此编码端可以通过第一目标虚拟扬声器的位置信息确定第一目标虚拟扬声器的 HOA 系数。若属性信息包括 HOA 系数时，编码端通过解码第一目标虚拟扬声器的属性信息可以获取到第一目标虚拟扬声器的 HOA 系数。

在一种可能的实现方式中，所述方法还包括：

根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

对所述第二虚拟扬声器信号进行编码，并写入所述码流。

在上述方案中，第二目标虚拟扬声器是编码端选择出的与第一目标虚拟编码器不相同的另一个目标虚拟扬声器。第一场景音频信号是待编码的原始场景音频信号，该第二目标虚拟扬声器可以是虚拟扬声器集合中的某一个虚拟扬声器，例如可以使用预先配置的目标虚拟扬声器选择策略从预设的虚拟扬声器集合中选择出第二目标虚拟扬声器。目标虚拟扬声器选择策略是从虚拟扬声器集合中选择与第一场景音频信号匹配的目标虚拟扬声器的策略，例如按照每个虚拟扬声器从第一场景音频信号中获取的声场成分来选择第二目标虚拟

扬声器。

在一种可能的实现方式中，所述方法还包括：

对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应地，所述对所述第二虚拟扬声器信号进行编码包括：

对所述对齐后的第二虚拟扬声器信号进行编码；

相应地，所述对所述第一虚拟扬声器信号进行编码，包括：

对所述对齐后的第一虚拟扬声器信号进行编码。

在上述方案中，编码端获取到对齐后的第一虚拟扬声器信号之后，可以对对齐后的第一虚拟扬声器信号进行编码，本申请实施例中通过将第一虚拟扬声器信号的各声道间重新调整对齐，增强了声道间相关性，有利于核心编码器对第一虚拟扬声器信号的编码处理。

在一种可能的实现方式中，所述方法还包括：

根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

相应地，所述对所述第一虚拟扬声器信号进行编码，包括：

根据所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得下混信号和边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；

对所述下混信号以及所述边信息进行编码。

在上述方案中，编码端在获取到第一虚拟扬声器信号和第二虚拟扬声器信号之后，编码端还可以根据第一虚拟扬声器信号和第二虚拟扬声器信号进行下混处理，以生成下混信号，例如对第一虚拟扬声器信号和第二虚拟扬声器信号进行幅度上的下混处理，以得到下混信号。另外还可以根据第一虚拟扬声器信号和第二虚拟扬声器信号生成边信息，边信息用于指示第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系，该关系具有多种实现方式，该边信息可以用于解码端针对下混信号进行上混，以恢复出第一虚拟扬声器信号和第二虚拟扬声器信号。例如边信息包括信号信息丢失分析参数，以使得解码端通过信号信息丢失分析参数恢复出第一虚拟扬声器信号和第二虚拟扬声器信号。

在一种可能的实现方式中，所述方法还包括：

对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应的，所述根据所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得下混信号和边信息，包括：

根据所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号获得所述下混信号和所述边信息；

相应的，所述边信息用于指示所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号之间的关系。

在上述方案中，编码端在生成下混信号之前，可以先执行虚拟扬声器信号的对齐操作，在完成对齐操作之后，再生成下混信号和边信息。本申请实施例中通过将第一虚拟扬声器

信号和第二虚拟扬声器的各声道间重新调整对齐，增强了声道间相关性，有利于核心编码器对第一虚拟扬声器信号的编码处理。

在一种可能的实现方式中，在根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器前，所述方法还包括：

根据编码速率和/或所述当前场景音频信号的信号类型信息确定是否需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器；

若需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器，才根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器。

在上述方案中，编码端还可以进行信号选择，以确定是否需要获取第二目标虚拟扬声器，在需要获取第二目标虚拟扬声器的情况下，编码端可以生成第二虚拟扬声器信号，在不需要获取第二目标虚拟扬声器的情况下，编码端可以不生成第二虚拟扬声器信号。其中，编码器可以根据音频编码器的配置信息和/或第一场景音频信号的信号类型信息进行决策，以确定在选择出第一目标虚拟扬声器之外是否还需要选择别的目标虚拟扬声器。例如，若编码速率高于预设的阈值，则确定需要获取两个主要声场成分对应的目标虚拟扬声器，则在确定出第一目标虚拟扬声器之外，还可以继续确定第二目标虚拟扬声器。又如，根据第一场景音频信号的信号类型信息确定需要获取包含声源方向占优的两个主要声场成分对应的目标虚拟扬声器，则在确定出第一目标虚拟扬声器之外，还可以继续确定第二目标虚拟扬声器。相反的，若根据编码速率和/或第一场景音频信号的信号类型信息确定只需要获取一个目标虚拟扬声器，则在确定第一目标虚拟扬声器之后，就确定不再获取除第一目标虚拟扬声器以外的目标虚拟扬声器。本申请实施例中通过信号选择，可以减少编码端进行编码的数据量，提高编码效率。

第二方面，本申请实施例还提供一种音频解码方法，包括：

接收码流；

解码所述码流以获得虚拟扬声器信号；

根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号。

在本申请实施例中，首先接收码流，然后解码码流以获得虚拟扬声器信号，最后根据目标虚拟扬声器的属性信息以及虚拟扬声器信号获得重建的场景音频信号。本申请实施例中可以从码流中解码得到虚拟扬声器信号，通过目标虚拟扬声器的属性信息和虚拟扬声器信号得到了重建的场景音频信号，本申请实施例中，获取到的码流中携带虚拟扬声器信号和残差信号，减少了解码的数据量，提高了解码效率。

在一种可能的实现方式中，所述方法还包括：

解码所述码流以获得所述目标虚拟扬声器的属性信息。

在上述方案中，编码端除了对虚拟扬声器进行编码，还可以对目标虚拟扬声器的属性信息进行编码，并将编码后的目标虚拟扬声器的属性信息写入到码流中，例如可以通过码流获取到第一目标虚拟扬声器的属性信息。本申请实施例中码流中可以携带编码后的第一目标虚拟扬声器的属性信息，使得解码端通过解码码流，就可以确定出第一目标虚拟扬声器的属性信息，便于解码端的音频解码。

在一种可能的实现方式中，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器

的高阶立体混响 HOA 系数；

所述根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号，包括：

对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

在上述方案中，解码端首先确定目标虚拟扬声器的 HOA 系数，例如解码端中可以预先存储目标虚拟扬声器的 HOA 系数，解码端获取到虚拟扬声器信号和目标虚拟扬声器的 HOA 系数之后，根据虚拟扬声器信号和目标虚拟扬声器的 HOA 系数可以得到重建的场景音频信号。从而提高重建的场景音频信号的质量。

在一种可能的实现方式中，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的位置信息；

所述根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号，包括：

根据所述目标虚拟扬声器的位置信息确定所述目标虚拟扬声器的 HOA 系数；

对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

在上述方案中，目标虚拟扬声器的属性信息可以包括：目标虚拟扬声器的位置信息。解码端预先存储虚拟扬声器集合中每个虚拟扬声器的 HOA 系数，解码端还存储有每个虚拟扬声器的位置信息，例如解码端可以根据虚拟扬声器的位置信息和该虚拟扬声器的 HOA 系数之间的对应关系确定出目标虚拟扬声器的位置信息对应的 HOA 系数，或者解码端可以根据目标虚拟扬声器的位置信息计算出目标虚拟扬声器的 HOA 系数。因此解码端可以通过目标虚拟扬声器的位置信息确定目标虚拟扬声器的 HOA 系数。解决了解码端需要确定目标虚拟扬声器的 HOA 系数的问题。

在一种可能的实现方式中，所述虚拟扬声器信号是根据第一虚拟扬声器信号和第二虚拟扬声器信号下混获得的下混信号，所述方法还包括：

解码所述码流以获得边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；

根据所述边信息和所述下混信号获得所述第一虚拟扬声器信号和所述第二虚拟扬声器信号；

相应的，所述根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号，包括：

根据所述目标虚拟扬声器的属性信息、所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得所述重建的场景音频信号。

在上述方案中，编码端根据第一虚拟扬声器信号和第二虚拟扬声器信号进行下混处理时生成下混信号，编码端还可以针对下混信号进行信号补偿，以生成边信息，该边信息可以被写入码流中，解码端可以通过码流得到边信息，解码端可以根据边信息进行信号补偿，以得到第一虚拟扬声器信号和第二虚拟扬声器信号，因此在进行信号重建时，可以使用第一虚拟扬声器信号和第二虚拟扬声器信号，以及前述的目标虚拟扬声器的属性信息，从而

提高解码端的解码信号质量。

第三方面，本申请实施例提供一种音频编码装置，包括：

获取模块，用于根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器；

信号生成模块，用于根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号；

编码模块，用于对所述第一虚拟扬声器信号进行编码，以得到码流。

在一种可能的实现方式中，所述获取模块，用于根据所述虚拟扬声器集合从所述当前场景音频信号中获取主要声场成分；根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器。

在本申请的第三方面中，音频编码装置的组成模块还可以执行前述第一方面以及各种可能的实现方式中所描述的步骤，详见前述对第一方面以及各种可能的实现方式中的说明。

在一种可能的实现方式中，所述获取模块，用于根据所述主要声场成分从高阶立体混响 HOA 系数集合中选择出与所述主要声场成分对应的 HOA 系数，所述 HOA 系数集合中的 HOA 系数与所述虚拟扬声器集合中的虚拟扬声器一一对应；确定所述虚拟扬声器集合中与所述主要声场成分对应的 HOA 系数对应的虚拟扬声器为所述第一目标虚拟扬声器。

在一种可能的实现方式中，所述获取模块，用于根据所述主要声场成分获取所述第一目标虚拟扬声器的配置参数；根据所述第一目标虚拟扬声器的配置参数生成所述第一目标虚拟扬声器对应的 HOA 系数；确定所述虚拟扬声器集合中所述第一目标虚拟扬声器对应的 HOA 系数对应的虚拟扬声器为所述目标虚拟扬声器。

在一种可能的实现方式中，所述获取模块，用于根据音频编码器的配置信息确定所述虚拟扬声器集合中的多个虚拟扬声器的配置参数；根据所述主要声场成分从所述多个虚拟扬声器的配置参数中选择出所述第一目标虚拟扬声器的配置参数。

在一种可能的实现方式中，所述第一目标虚拟扬声器的配置参数包括：所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息；

所述获取模块，用于根据所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息确定所述第一目标虚拟扬声器对应的 HOA 系数。

在一种可能的实现方式中，所述编码模块，还用于对所述第一目标虚拟扬声器的属性信息进行编码，并写入所述码流。

在一种可能的实现方式中，所述当前场景音频信号，包括：待编码 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的 HOA 系数；

所述信号生成模块，用于对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

在一种可能的实现方式中，所述当前场景音频信号包括：待编码高阶立体混响 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的位置信息；

所述信号生成模块，用于根据所述第一目标虚拟扬声器的位置信息获取所述第一目标虚拟扬声器对应的 HOA 系数；对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

在一种可能的实现方式中，所述获取模块，用于根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

所述信号生成模块，用于根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

所述编码模块，用于对所述第二虚拟扬声器信号进行编码，并写入所述码流。

在一种可能的实现方式中，所述信号生成模块，用于对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应地，所述编码模块，用于对所述对齐后的第二虚拟扬声器信号进行编码；

相应地，所述编码模块，用于对所述对齐后的第一虚拟扬声器信号进行编码。

在一种可能的实现方式中，所述获取模块，用于根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

所述信号生成模块，用于根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

相应地，所述编码模块，用于根据所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得下混信号和边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；对所述下混信号以及所述边信息进行编码。

在一种可能的实现方式中，所述信号生成模块，用于对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应的，所述编码模块，用于根据所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号获得所述下混信号和所述边信息；

相应的，所述边信息用于指示所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号之间的关系。

在一种可能的实现方式中，所述获取模块，用于在根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器前，根据编码速率和/或所述当前场景音频信号的信号类型信息确定是否需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器；若需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器，才根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器。

第四方面，本申请实施例提供一种音频解码装置，包括：

接收模块，用于接收码流；

解码模块，用于解码所述码流以获得虚拟扬声器信号；

重建模块，用于根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号。

在一种可能的实现方式中，所述解码模块，还用于解码所述码流以获得所述目标虚拟扬声器的属性信息。

在一种可能的实现方式中，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的高阶立体混响 HOA 系数；

所述重建模块，用于对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

在一种可能的实现方式中，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的位置信息；

所述重建模块，用于根据所述目标虚拟扬声器的位置信息确定所述目标虚拟扬声器的 HOA 系数；对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

在一种可能的实现方式中，所述虚拟扬声器信号是根据第一虚拟扬声器信号和第二虚拟扬声器信号下混获得的下混信号，所述装置还包括：信号补偿模块，其中，

所述解码模块，用于解码所述码流以获得边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；

所述信号补偿模块，用于根据所述边信息和所述下混信号获得所述第一虚拟扬声器信号和所述第二虚拟扬声器信号；

相应的，所述重建模块，用于根据所述目标虚拟扬声器的属性信息、所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得所述重建的场景音频信号。

在本申请的第四方面中，音频解码装置的组成模块还可以执行前述第二方面以及各种可能的实现方式中所描述的步骤，详见前述对第二方面以及各种可能的实现方式中的说明。

第五方面，本申请实施例提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有指令，当其在计算机上运行时，使得计算机执行上述第一方面或第二方面所述的方法。

第六方面，本申请实施例提供了一种包含指令的计算机程序产品，当其在计算机上运行时，使得计算机执行上述第一方面或第二方面所述的方法。

第七方面，本申请实施例提供一种通信装置，该通信装置可以包括终端设备或者芯片等实体，所述通信装置包括：处理器，可选的，所述通信装置还包括存储器；所述存储器用于存储指令；所述处理器用于执行所述存储器中的所述指令，使得所述通信装置执行如前述第一方面或第二方面中任一项所述的方法。

第八方面，本申请提供了一种芯片系统，该芯片系统包括处理器，用于支持音频编码装置或者音频解码装置实现上述方面中所涉及的功能，例如，发送或处理上述方法中所涉及的数据和/或信息。在一种可能的设计中，所述芯片系统还包括存储器，所述存储器，用于保存音频编码装置或者音频解码装置必要的程序指令和数据。该芯片系统，可以由芯片构成，也可以包括芯片和其他分立器件。

第九方面，本申请提供一种计算机可读存储介质，包括如前述第一方面中任一项所述的方法所生成的码流。

附图说明

图 1 为本申请实施例提供的音频处理系统的组成结构示意图；

图 2a 为本申请实施例提供的音频编码器和音频解码器应用于终端设备的示意图；

图 2b 为本申请实施例提供的音频编码器应用于无线设备或者核心网设备的示意图；

图 2c 为本申请实施例提供的音频解码器应用于无线设备或者核心网设备的示意图；
图 3a 为本申请实施例提供的多声道编码器和多声道解码器应用于终端设备的示意图；
图 3b 为本申请实施例提供的多声道编码器应用于无线设备或者核心网设备的示意图；
图 3c 为本申请实施例提供的多声道解码器应用于无线设备或者核心网设备的示意图；
图 4 为本申请实施例中音频编码装置和音频解码装置之间的一种交互流程示意图；
图 5 为本申请实施例提供的编码端的一种结构示意图；
图 6 为本申请实施例提供的解码端的一种结构示意图；
图 7 为本申请实施例提供的编码端的一种结构示意图；
图 8 为本申请实施例提供的一种球面上近似均匀分布的虚拟扬声器的示意图；
图 9 为本申请实施例提供的编码端的一种结构示意图；
图 10 为本申请实施例提供的一种音频编码装置的组成结构示意图；
图 11 为本申请实施例提供的一种音频解码装置的组成结构示意图；
图 12 为本申请实施例提供的另一种音频编码装置的组成结构示意图；
图 13 为本申请实施例提供的另一种音频解码装置的组成结构示意图。

具体实施方式

本申请实施例提供了一种音频编解码方法和装置，用于减少编码场景音频信号的数据量，提高编解码效率。

下面结合附图，对本申请的实施例进行描述。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的术语在适当情况下可以互换，这仅仅是描述本申请的实施例中对相同属性的对象在描述时所采用的区分方式。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，以便包含一系列单元的过程、方法、系统、产品或设备不必限于那些单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它单元。

本申请实施例的技术方案可以应用于各种的音频处理系统，如图 1 所示，为本申请实施例提供的音频处理系统的组成结构示意图。音频处理系统 100 可以包括：音频编码装置 101 和音频解码装置 102。其中，音频编码装置 101 可用于生成码流，然后该音频编码码流可以通过音频传输通道传输给音频解码装置 102，音频解码装置 102 可以接收到码流，然后执行音频解码装置 102 的音频解码功能，最后得到重建后的信号。

在本申请的实施例中，该音频编码装置可以应用于各种有音频通信需要的终端设备、有转码需要的无线设备与核心网设备，例如音频编码装置可以是上述终端设备或者无线设备或者核心网设备的音频编码器。同样的，该音频解码装置可以应用于各种有音频通信需要的终端设备、有转码需要的无线设备与核心网设备，例如音频解码装置可以是上述终端设备或者无线设备或者核心网设备的音频解码器。例如，音频编码器可以包括无线接入网、核心网的媒体网关、转码设备、媒体资源服务器、移动终端、固网终端等，音频编码器还可以是应用于虚拟现实技术(virtual reality，VR）流媒体（streaming）服务中的音频编解码器。

在申请实施例中，以适用于虚拟现实流媒体（VR streaming）服务中的音频编解码模块（audio encoding及audio decoding）为例，端到端对音频信号的处理流程包括：音频信号A经过采集模块（acquisition）后进行预处理操作（audio preprocessing），预处理操作包括滤除掉信号中的低频部分，可以是以20Hz或者50Hz为分界点，提取信号中的方位信息，之后进行编码处理（audio encoding）打包（file/segment encapsulation）之后发送(delivery)到解码端，解码端首先进行解包（file/segment decapsulation），之后解码（audio decoding），对解码信号进行双耳渲染（audio rendering）处理，渲染处理后的信号映射到收听者耳机（headphones）上，可以为独立的耳机也可以是眼镜设备上的耳机。

如图2a所示，为本申请实施例提供的音频编码器和音频解码器应用于终端设备的示意图。对于每个终端设备都可以包括：音频编码器、信道编码器、音频解码器、信道解码器。具体的，信道编码器用于对音频信号进行信道编码，信道解码器用于对音频信号进行信道解码。例如，在第一终端设备20中可以包括：第一音频编码器201、第一信道编码器202、第一音频解码器203、第一信道解码器204。在第二终端设备21中可以包括：第二音频解码器211、第二信道解码器212、第二音频编码器213、第二信道编码器214。第一终端设备20连接无线或者有线的第一网络通信设备22，第一网络通信设备22和无线或者有线的第二网络通信设备23之间通过数字信道连接，第二终端设备21连接无线或者有线的第二网络通信设备23。其中，上述无线或者有线的网络通信设备可以泛指信号传输设备，例如通信基站，数据交换设备等。

在音频通信中，作为发送端的终端设备首先进行音频采集，对采集到的音频信号进行音频编码，再进行信道编码后，通过无线网络或者核心网进行在数字信道中传输。而作为接收端的终端设备根据接收到的信号进行信道解码，以得到码流，然后经过音频解码恢复出音频信号，由接收端的终端设备进音频回放。

如图2b所示，为本申请实施例提供的音频编码器应用于无线设备或者核心网设备的示意图。其中，无线设备或者核心网设备25包括：信道解码器251、其他音频解码器252、本申请实施例提供的音频编码器253、信道编码器254，其中，其他音频解码器252是指除音频解码器以外的其他音频解码器。在无线设备或者核心网设备25内，首先通过信道解码器251对进入该设备的信号进行信道解码，然后使用其他音频解码器252进行音频解码，然后使用本申请实施例提供的音频编码器253进行音频编码，最后使用信道编码器254对音频信号进行信道编码，完成信道编码之后再传输出去。其中，其他音频解码器252是对信道解码器251解码后的码流进行音频解码。

如图2c所示，为本申请实施例提供的音频解码器应用于无线设备或者核心网设备的示意图。其中，无线设备或者核心网设备25包括：信道解码器251、本申请实施例提供的音频解码器255、其他音频编码器256、信道编码器254，其中，其他音频编码器256是指除音频编码器以外的其他音频编码器。在无线设备或者核心网设备25内，首先通过信道解码器251对进入该设备的信号进行信道解码，然后使用音频解码器255对接收到的音频编码码流进行解码，然后使用其他音频编码器256进行音频编码，最后使用信道编码器254对音频信号进行信道编码，完成信道编码之后再传输出去。在无线设备或者核心网设备中，

如果需要实现转码，则需要进行相应的音频编解码处理。其中，无线设备指的是通信中的射频相关的设备，核心网设备指的是通信中核心网相关的设备。

在本申请的一些实施例中，该音频编码装置可以应用于各种有音频通信需要的终端设备、有转码需要的无线设备与核心网设备，例如音频编码装置可以是上述终端设备或者无线设备或者核心网设备的多声道编码器。同样的，该音频解码装置可以应用于各种有音频通信需要的终端设备、有转码需要的无线设备与核心网设备，例如音频解码装置可以是上述终端设备或者无线设备或者核心网设备的多声道解码器。

如图 3a 所示，为本申请实施例提供的多声道编码器和多声道解码器应用于终端设备的示意图，对于每个终端设备都可以包括：多声道编码器、信道编码器、多声道解码器、信道解码器。该多声道编码器可以执行本申请实施例提供的音频编码方法，该多声道解码器可以执行本申请实施例提供的音频解码方法。具体的，信道编码器用于对多声道信号进行信道编码，信道解码器用于对多声道信号进行信道解码。例如，在第一终端设备 30 中可以包括：第一多声道编码器 301、第一信道编码器 302、第一多声道解码器 303、第一信道解码器 304。在第二终端设备 31 中可以包括：第二多声道解码器 311、第二信道解码器 312、第二多声道编码器 313、第二信道编码器 314。第一终端设备 30 连接无线或者有线的第一网络通信设备 32，第一网络通信设备 32 和无线或者有线的第二网络通信设备 33 之间通过数字信道连接，第二终端设备 31 连接无线或者有线的第二网络通信设备 33。其中，上述无线或者有线的网络通信设备可以泛指信号传输设备，例如通信基站，数据交换设备等。音频通信中作为发送端的终端设备对采集到的多声道信号进行多声道编码，再进行信道编码后，通过无线网络或者核心网进行在数字信道中传输。而作为接收端的终端设备根据接收到的信号，进行信道解码，以得到多声道信号编码码流，然后经过多声道解码恢复出多声道信号，由作为接收端的终端设备进回放。

如图 3b 所示，为本申请实施例提供的多声道编码器应用于无线设备或者核心网设备的示意图，其中，无线设备或者核心网设备 35 包括：信道解码器 351、其他音频解码器 352、多声道编码器 353、信道编码器 354，与前述图 2b 类似，此处不再赘述。

如图 3c 所示，为本申请实施例提供的多声道解码器应用于无线设备或者核心网设备的示意图，其中，无线设备或者核心网设备 35 包括：信道解码器 351、多声道解码器 355、其他音频编码器 356、信道编码器 354，与前述图 2c 类似，此处不再赘述。

其中，音频编码处理可以是多声道编码器中的一部分，音频解码处理可以是多声道解码器中的一部分，例如，对采集到的多声道信号进行多声道编码可以是将采集到的多声道信号经过处理后得到音频信号，再按照本申请实施例提供的方法对得到的音频信号进行编码；解码端根据多声道信号编码码流，解码得到音频信号，经过上混处理后恢复出多声道信号。因此，本申请实施例也可应用于终端设备、无线设备、核心网设备中的多声道编码器和多声道解码器。在无线或者核心网设备中，如果需要实现转码，则需要进行相应的多声道编解码处理。

本申请实施例提供的音频编解码方法可以包括：音频编码方法和音频解码方法，其中，音频编码方法由音频编码装置执行，音频解码方法由音频解码装置执行，音频编码装置和音频解码装置之间可以进行通信。接下来基于前述的系统架构以及音频编码装置和音频解

码装置，对本申请实施例提供的音频编码方法和音频解码方法进行说明。如图 4 所示，为本申请实施例中音频编码装置和音频解码装置之间的一种交互流程示意图，其中，下述步骤 401 至步骤 403 可以由音频编码装置（如下简称编码端）执行，下述步骤 411 至步骤 413 可以由音频解码装置（如下简称解码端）执行，主要包括如下过程：

401、根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器。

其中，编码端获取当前场景音频信号，该当前场景音频信号是指对空间中麦克风所在位置的声场进行采集得到的音频信号，当前场景音频信号也可以称为原始场景音频信号。例如当前场景音频信号可以是通过高阶立体混响(higher order ambisonics，HOA) 技术得到的音频信号。

本申请实施例中，编码端可以预先配置虚拟扬声器集合，该虚拟扬声器集合中可以包括多个虚拟扬声器，场景音频信号在实际回放时，可以通过耳机回放，也可以通过布置在房间中的多个扬声器回放。使用扬声器回放时，基本方法是通过多个扬声器的信号进行叠加，使得空间中某点（听音人所在的位置）声场在某个标准下尽可能的接近录制场景音频信号时的原始声场。本申请实施例中使用虚拟扬声器计算场景音频信号对应的回放信号，使用该回放信号作为传输信号，并进而生成压缩后的信号。虚拟扬声器表示的是在空间声场中虚拟存在的扬声器，该虚拟扬声器可以实现在编码端的场景音频信号的回放。

本申请实施例中，虚拟扬声器集合中包括多个虚拟扬声器，这多个虚拟扬声器中每个虚拟扬声器对应有虚拟扬声器配置参数（简称配置参数）。虚拟扬声器配置参数包括且不限于：虚拟扬声器的个数，虚拟扬声器的 HOA 阶数、虚拟扬声器的位置坐标等信息。编码端获取到上述的虚拟扬声器集合之后，根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器，当前场景音频信号是待编码的原始场景音频信号，该第一目标虚拟扬声器可以是虚拟扬声器集合中的某一个虚拟扬声器，例如可以使用预先配置的目标虚拟扬声器选择策略从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器。目标虚拟扬声器选择策略是从虚拟扬声器集合中选择与当前场景音频信号匹配的目标虚拟扬声器的策略，例如按照每个虚拟扬声器从当前场景音频信号中获取的声场成分来选择第一目标虚拟扬声器。又如，按照每个虚拟扬声器的位置信息从当前场景音频信号中选择第一目标虚拟扬声器。其中，第一目标虚拟扬声器为虚拟扬声器集合中用于回放当前场景音频信号的虚拟扬声器，即编码端可以从虚拟扬声器集合中选择出可回放当前场景音频信号的目标虚拟编码器。

不限定的是，本申请实施例中，通过步骤 401 选择出第一目标虚拟扬声器之后，可以执行后续针对第一目标虚拟扬声器的处理过程，如后续步骤 402 至步骤 403。本申请实施例不仅可以选择出第一目标虚拟扬声器，还可以选择出更多的目标虚拟扬声器，例如还可以选择出第二目标虚拟扬声器，针对第二目标虚拟扬声器，同样需要执行与后续步骤 402 至 403 相类似的过程，详见后续实施例的说明。

在本申请实施例中，编码端选择出第一目标虚拟扬声器之后，编码端还可以获取第一目标虚拟扬声器的属性信息，第一目标虚拟扬声器的属性信息包括与第一目标虚拟扬声器的属性相关的信息，该属性信息可以根据具体应用场景设置，例如第一目标虚拟扬声器的属性信息包括：该第一目标虚拟扬声器的位置信息，或者该第一目标虚拟扬声器的 HOA 系

数。其中，第一目标虚拟扬声器的位置信息可以是该第一目标虚拟扬声器在空间的分布位置，也可以是该第一目标虚拟扬声器在虚拟扬声器集合中相对于其它虚拟扬声器的位置的信息，具体此处不做限定。虚拟扬声器集合中每个虚拟扬声器都对应有 HOA 系数，该 HOA 系数也可以称为 Ambisonic 系数，接下来对虚拟扬声器对应的 HOA 系数进行说明。

例如，HOA 阶数可以为 2 阶至 10 阶中的其中 1 个阶数，录制音频信号时的信号采样率为 48 至 192 千赫兹（kHz），采样深度为 16 或者 24 比特（bit），通过虚拟扬声器的 HOA 系数和场景音频信号可以生成 HOA 信号，HOA 信号的特点是带有声场的空间信息，HOA 信号是描述空间某点声场信号一定精度的信息。因此，可以考虑使用另一种表示形式描述某一位置点的声场信号，这种描述方法能够使用更少的数据量对空间位置点的信号达到同样精确度的描述，从而能达到信号压缩的目的。空间声场可以分解为多个平面波的叠加。因此，理论上可以将 HOA 信号表达的声场，重新使用多个平面波的叠加来表达，每个平面波使用一个声道的音频信号和一个方向向量表示。平面波叠加的表示形式能够使用更少的声道数目准确的表达原始声场，以达到信号压缩的目的。

在本申请的一些实施例中，编码端除了执行前述步骤 401，本申请实施例提供的音频编码方法还包括如下步骤：

A1、根据虚拟扬声器集合从当前场景音频信号中获取主要声场成分。

其中，步骤 A1 中的主要声场成分也可以称为第一主要声场成分。

在执行步骤 A1 的场景下，前述步骤 401 根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器，包括：

B1、根据主要声场成分从虚拟扬声器集合中选择出第一目标虚拟扬声器。

其中，编码端获取虚拟扬声器集合，编码端使用该虚拟扬声器集合对当前场景音频信号进行信号分解，以得到当前场景音频信号对应的主要声场成分。其中，主要声场成分表示的是当前场景音频信号中的主要声场所对应的音频信号。例如虚拟扬声器集合中包括多个虚拟扬声器，根据多个虚拟扬声器可以从当前场景音频信号中获取多个声场成分，即每个虚拟扬声器可以从当前场景音频信号中获取一个声场成分，接下来从多个声场成分中选择出主要声场成分，例如主要声场成分可以是多个声场成分中取值最大的一个或几个声场成分，或主要声场成分可以是多个声场成分中方向占优的一个或几个声场成分。虚拟扬声器集合中的每个虚拟扬声器对应一个声场成分，则根据主要声场成分从虚拟扬声器集合中选择出第一目标虚拟扬声器，例如主要声场成分对应的虚拟扬声器就是编码端选择出的第一目标虚拟扬声器。本申请实施例中，编码端通过主要声场成分可以选择出第一目标虚拟扬声器，解决了编码端需要确定第一目标虚拟扬声器的问题。

不限定的是，本申请实施例中，编码端具有多种方式选择出第一目标虚拟扬声器，例如编码端可以预设指定位置的虚拟扬声器作为第一目标虚拟扬声器，即按照虚拟扬声器集合中每个虚拟扬声器的位置选择出符合指定位置的虚拟扬声器作为第一目标虚拟扬声器。

其中，在本申请的一些实施例中，前述步骤 B1 根据主要声场成分从虚拟扬声器集合中选择出第一目标虚拟扬声器，包括：

根据主要声场成分从高阶立体混响 HOA 系数集合中选择出与主要声场成分对应的 HOA 系数，HOA 系数集合中的 HOA 系数与虚拟扬声器集合中的虚拟扬声器一一对应；

确定虚拟扬声器集合中与主要声场成分对应的 HOA 系数对应的虚拟扬声器为第一目标虚拟扬声器。

其中，编码端中根据虚拟扬声器集合预先配置 HOA 系数集合，HOA 系数集合中的 HOA 系数与虚拟扬声器集合中的虚拟扬声器之间的一一对应关系，因此根据主要声场成分选择出 HOA 系数之后，再根据上述一一对应关系从虚拟扬声器集合中查找与主要声场成分对应的 HOA 系数对应的目标虚拟扬声器，该查找出的目标虚拟扬声器即为第一目标虚拟扬声器，解决了编码端需要确定第一目标虚拟扬声器的问题。举例说明如下，HOA 系数集合中包括 HOA 系数 1、HOA 系数 2、HOA 系数 3，虚拟扬声器集合中包括虚拟扬声器 1、虚拟扬声器 2、虚拟扬声器 3，其中，HOA 系数集合中的 HOA 系数与虚拟扬声器集合中的虚拟扬声器一一对应，例如：HOA 系数 1 与虚拟扬声器 1 对应，HOA 系数 2 与虚拟扬声器 2 对应，HOA 系数 3 与虚拟扬声器 3 对应。若根据主要声场成分从 HOA 系数集合中选择出 HOA 系数 3，则可以确定第一目标虚拟扬声器为虚拟扬声器 3。

其中，在本申请的一些实施例中，前述步骤 B1 根据主要声场成分从虚拟扬声器集合中选择出第一目标虚拟扬声器，还包括：

C1、根据主要声场成分获取第一目标虚拟扬声器的配置参数；

C2、根据第一目标虚拟扬声器的配置参数生成第一目标虚拟扬声器对应的 HOA 系数；

C3、确定虚拟扬声器集合中第一目标虚拟扬声器对应的 HOA 系数对应的虚拟扬声器为第一目标虚拟扬声器。

其中，编码端在获取到主要声场成分之后，可以根据该主要声场成分可用于确定第一目标虚拟扬声器的配置参数，例如主要声场成分是多个声场成分中取值最大的一个或几个声场成分，或主要声场成分可以是多个声场成分中方向占优的一个或几个声场成分，该主要声场成分可用于确定出当前场景音频信号匹配的第一目标虚拟扬声器，第一目标虚拟扬声器配置有相应的属性信息，使用第一目标虚拟扬声器的配置参数都可以生成该第一目标虚拟扬声器的 HOA 系数，HOA 系数的生成过程可以通过 HOA 算法来实现，此处不再详细说明。虚拟扬声器集合中每个虚拟扬声器都对应有 HOA 系数，因此可以根据每个虚拟扬声器对应的 HOA 系数从虚拟扬声器集合中选择出第一目标虚拟扬声器，解决了编码端需要确定第一目标虚拟扬声器的问题。

其中，在本申请的一些实施例中，步骤 C1 根据主要声场成分获取第一目标虚拟扬声器的配置参数，包括：

根据音频编码器的配置信息确定虚拟扬声器集合中的多个虚拟扬声器的配置参数；

根据主要声场成分从多个虚拟扬声器的配置参数中选择出第一目标虚拟扬声器的配置参数。

其中，音频编码器中可以预先存储多个虚拟扬声器各自的配置参数，每个虚拟扬声器的配置参数可以是通过音频编码器的配置信息确定，音频编码器是指前述的编码端，该音频编码器的配置信息，包括且不限于：HOA 阶数、编码比特率等。音频编码器的配置信息可以用于确定虚拟扬声器的个数，以及每个虚拟扬声器的位置参数，解决了编码端需要确定虚拟扬声器的配置参数的问题。举例说明如下，若编码比特率较低时可以配置较少数量的虚拟扬声器，若编码比特率较高时可以配置多个数量的虚拟扬声器。又如虚拟扬声器的

HOA 阶数可以等于音频编码器的 HOA 阶数。不限定的是，本申请实施例中，除了通过音频编码器的配置信息确定多个虚拟扬声器各自的配置参数之外，还可以根据用户自定义信息多个虚拟扬声器各自的配置参数，例如，用户可以自定义虚拟扬声器的位置、HOA 阶数、虚拟扬声器的个数等。

编码端从虚拟扬声器集合中获取多个虚拟扬声器的配置参数，对于每个虚拟扬声器而言，都存在相应的虚拟扬声器配置参数，每个虚拟扬声器配置参数包括且不限于：虚拟扬声器的 HOA 阶数、虚拟扬声器的位置坐标等信息。使用每个虚拟扬声器的配置参数都可以生成该虚拟扬声器的 HOA 系数，HOA 系数的生成过程可以通过 HOA 算法来实现，此处不再详细说明。针对虚拟扬声器集合中的每个虚拟扬声器分别生成一个 HOA 系数，虚拟扬声器集合中所有虚拟扬声器分别配置的 HOA 系数构成 HOA 系数集合，解决了编码端需要确定虚拟扬声器集合中每个虚拟扬声器的 HOA 系数的问题。

其中，在本申请的一些实施例中，第一目标虚拟扬声器的配置参数包括：第一目标虚拟扬声器的位置信息和 HOA 阶数信息；

前述步骤 C2 根据第一目标虚拟扬声器的配置参数生成第一目标虚拟扬声器对应的 HOA 系数，包括：

根据第一目标虚拟扬声器的位置信息和 HOA 阶数信息确定第一目标虚拟扬声器对应的 HOA 系数。

其中，虚拟扬声器集合中的每个虚拟扬声器的配置参数都可以包括该虚拟扬声器的位置信息以及该虚拟扬声器的 HOA 阶数信息。同样的，第一目标虚拟扬声器的配置参数包括：第一目标虚拟扬声器的位置信息和 HOA 阶数信息。例如可以按照局部等距的虚拟扬声器空间分布方式确定虚拟扬声器集合中每个虚拟扬声器的位置信息，局部等距的虚拟扬声器空间分布方式是指多个虚拟扬声器在空间中按照局部等距的方式进行分布，例如局部等距可以包括：均匀分布或者不均匀分布。使用每个虚拟扬声器的位置信息和 HOA 阶数信息都可以生成该虚拟扬声器的 HOA 系数，HOA 系数的生成过程可以通过 HOA 算法来实现，解决了编码端需要确定第一目标虚拟扬声器的 HOA 系数的问题。

另外，本申请实施例中针对虚拟扬声器集合中的每个虚拟扬声器分别生成一组 HOA 系数，多组 HOA 系数构成前述的 HOA 系数集合。虚拟扬声器集合中所有虚拟扬声器分别配置的 HOA 系数构成 HOA 系数集合，解决了编码端需要确定虚拟扬声器集合中每个虚拟扬声器的 HOA 系数的问题。

402、根据当前场景音频信号和第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号。

其中，编码端获取到当前场景音频信号和第一目标虚拟扬声器的属性信息之后，编码端可以进行当前场景音频信号的回放，编码端根据当前场景音频信号和第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，该第一虚拟扬声器信号即为当前场景音频信号的回放信号。第一目标虚拟扬声器的属性信息描述了与第一目标虚拟扬声器的属性相关的信息，该第一目标虚拟扬声器是编码端选择出的可回放当前场景音频信号的虚拟扬声器，因此通过第一目标虚拟扬声器的属性信息对当前场景音频信号进行回放，可以得到第一虚拟扬声器信号。该第一虚拟扬声器信号的数据量大小与当前场景音频信号的声道数无关，该

第一虚拟扬声器信号的数据量大小与第一目标虚拟扬声器有关。例如，本申请实施例中，第一虚拟扬声器信号相比于当前场景音频信号，采用较少的声道进行表示，例如当前场景音频信号为 3 阶 HOA 信号，该 HOA 信号为 16 个声道，本申请实施例中可以将 16 个声道压缩为 2 个声道，即编码端生成的虚拟扬声器信号为 2 个声道，例如编码端生成的虚拟扬声器信号可以包括前述的第一虚拟扬声器信号和第二虚拟扬声器信号等，编码端生成的虚拟扬声器信号的声道数与第一场景音频信号的声道数无关。通过后续步骤描述可知，码流中可以携带 2 个声道的第一虚拟扬声器信号，相应的，解码端接收该码流，解码该码流得到的虚拟扬声器信号为 2 个声道，解码端通过 2 个声道的虚拟扬声器信号可以重建出 16 个声道的场景音频信号，且保证了重建的场景音频信号与原始的场景音频信号相比时，具有主观和客观质量相当的效果。

可以理解的是，前述步骤 401 和步骤 402 具体可以由空间编码器来实现动态图像专家组（moving picture experts group，MPEG）空间编码器来实现。

在本申请的一些实施例中，当前场景音频信号可以包括：待编码 HOA 信号；第一目标虚拟扬声器的属性信息包括第一目标虚拟扬声器的 HOA 系数；

步骤 402 根据当前场景音频信号和第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，包括：

对待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数进行线性组合，以得到第一虚拟扬声器信号。

其中，以当前场景音频信号为待编码 HOA 信号为例，编码端首先确定第一目标虚拟扬声器的 HOA 系数，例如编码端根据主要声场成分从 HOA 系数集合中选择出 HOA 系数，该选择出的 HOA 系数就是第一目标虚拟扬声器的 HOA 系数，编码端获取到待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数之后，根据待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数可以生成第一虚拟扬声器信号，其中，待编码 HOA 信号可以采用第一目标虚拟扬声器的 HOA 系数进行线性组合得到，第一虚拟扬声器信号的求解可以被转换为对线性组合的求解问题。

例如，第一目标虚拟扬声器的属性信息可以包括：第一目标虚拟扬声器的 HOA 系数。编码端通过解码第一目标虚拟扬声器的属性信息可以获取到第一目标虚拟扬声器的 HOA 系数。编码端对待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数进行线性组合，即编码端将待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数组合在一起，可以得到线性组合矩阵，接下来编码端可以对线性组合矩阵进行求最优解，得到的最优解就是第一虚拟扬声器信号。其中，该最优解与对线性组合矩阵进行求解时采用的算法有关。本申请实施例解决了编码端需要生成第一虚拟扬声器信号的问题。

在本申请的一些实施例中，当前场景音频信号包括：待编码高阶立体混响 HOA 信号；第一目标虚拟扬声器的属性信息包括第一目标虚拟扬声器的位置信息；

步骤 402 根据当前场景音频信号和第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，包括：

根据第一目标虚拟扬声器的位置信息获取第一目标虚拟扬声器对应的 HOA 系数；

对待编码 HOA 信号和第一目标虚拟扬声器对应的 HOA 系数进行线性组合，以得到第一

虚拟扬声器信号。

其中，第一目标虚拟扬声器的属性信息可以包括：第一目标虚拟扬声器的位置信息，编码端预先存储虚拟扬声器集合中每个虚拟扬声器的 HOA 系数，编码端还存储有每个虚拟扬声器的位置信息，虚拟扬声器的位置信息和该虚拟扬声器的 HOA 系数之间存在对应关系，因此编码端可以通过第一目标虚拟扬声器的位置信息确定第一目标虚拟扬声器的 HOA 系数。若属性信息包括 HOA 系数时，编码端通过解码第一目标虚拟扬声器的属性信息可以获取到第一目标虚拟扬声器的 HOA 系数。

编码端获取待编码 HOA 信号以及第一目标虚拟扬声器的 HOA 系数之后，编码端对待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数进行线性组合，即编码端将待编码 HOA 信号和第一目标虚拟扬声器的 HOA 系数组合在一起，可以得到线性组合矩阵，接下来编码端可以对线性组合矩阵进行求最优解，得到的最优解就是第一虚拟扬声器信号。

举例说明如下，第一目标虚拟扬声器的 HOA 系数用矩阵 A 表示，用矩阵 A 可以线性组合出待编码 HOA 信号，其中可以采用最小二乘方法求得理论的最优解w，即为第一虚拟扬声器信号，例如可以采用如下计算式：

$$w = A^{-1}X,$$

其中，A^{-1} 代表矩阵 A 的逆矩阵，矩阵 A 的大小为$(M \times C)$，C 为第一目标虚拟扬声器个数，M 为N阶的 HOA 系数的声道个数，a 表示第一目标虚拟扬声器的 HOA 系数，例如，

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1C} \\ \vdots & & \vdots \\ a_{M1} & \cdots & a_{MC} \end{bmatrix}。$$

其中，X 代表待编码 HOA 信号，矩阵 X 的大小为(M×L)，M 为N阶的 HOA 系数的声道个数，L 为采样点数，x 表示待编码 HOA 信号的系数，例如，

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1L} \\ \vdots & & \vdots \\ x_{M1} & \cdots & x_{ML} \end{bmatrix}。$$

403、对虚拟扬声器信号进行编码，以得到码流。

本申请实施例中，编码端在生成第一虚拟扬声器信号之后，编码端可以对第一虚拟扬声器信号进行编码，以得到码流。例如编码端具体可以是核心编码器，核心编码器对第一虚拟扬声器信号进行编码，以得到码流。该码流也可以称为音频信号编码码流。本申请实施例编码端对该第一虚拟扬声器信号进行编码，而不再对场景音频信号进行编码，通过选择出的第一目标虚拟扬声器，使得空间中听音人所在的位置声场尽可能的接近录制场景音频信号时的原始声场，保证了编码端的编码质量，且第一虚拟扬声器信号的编码数据量与场景音频信号的声道数无关，减少编码场景音频信号的数据量，提高编解码效率。

在本申请的一些实施例中，编码端执行上述的步骤 401 至步骤 403 之后，本申请实施例提供的音频编码方法还包括如下步骤：

对第一目标虚拟扬声器的属性信息进行编码，并写入码流。

其中，编码端除了对虚拟扬声器进行编码，还可以对第一目标虚拟扬声器的属性信息进行编码，并将编码后的第一目标虚拟扬声器的属性信息写入到码流中，此时得到的码流中可以包括：编码后的虚拟扬声器和编码后的第一目标虚拟扬声器的属性信息。本申请实施例中码流中可以携带编码后的第一目标虚拟扬声器的属性信息，使得解码端通过解码码流，就可以确定出第一目标虚拟扬声器的属性信息，便于解码端的音频解码。

需要说明的是，前述步骤 401 至步骤 403 中描述了从虚拟扬声器集合中选择出第一目标扬声器的情况下，基于该第一目标虚拟扬声器生成第一虚拟扬声器信号，并根据第一虚拟扬声器进行信号编码的过程。不限定的是，本申请实施例中，编码端不仅可以选择出第一目标虚拟扬声器，还可以选择出更多的目标虚拟扬声器，例如还可以选择出第二目标虚拟扬声器，针对第二目标虚拟扬声器，同样需要执行与前述步骤 402 至 403 相类似的过程，接下来进行详细的说明。

在本申请的一些实施例中，编码端除了执行前述步骤之外，本申请实施例提供的音频编码方法还包括：

D1、根据第一场景音频信号从虚拟扬声器集合中选择出第二目标虚拟扬声器；

D2、根据第一场景音频信号和第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

D3、对第二虚拟扬声器信号进行编码，并写入码流。

其中，步骤 D1 的实现方式与前述步骤 401 相类似，第二目标虚拟扬声器是编码端选择出的与第一目标虚拟编码器不相同的另一个目标虚拟扬声器。第一场景音频信号是待编码的原始场景音频信号，该第二目标虚拟扬声器可以是虚拟扬声器集合中的某一个虚拟扬声器，例如可以使用预先配置的目标虚拟扬声器选择策略从预设的虚拟扬声器集合中选择出第二目标虚拟扬声器。目标虚拟扬声器选择策略是从虚拟扬声器集合中选择与第一场景音频信号匹配的目标虚拟扬声器的策略，例如按照每个虚拟扬声器从第一场景音频信号中获取的声场成分来选择第二目标虚拟扬声器。

在本申请的一些实施例中，本申请实施例提供的音频编码方法还包括如下步骤：

E1、根据虚拟扬声器集合从第一场景音频信号中获取第二主要声场成分。

在执行步骤 E1 的场景下，前述步骤 D1 根据第一场景音频信号从预设的虚拟扬声器集合中选择出第二目标虚拟扬声器，包括：

F1、根据第二主要声场成分从虚拟扬声器集合中选择出第二目标虚拟扬声器。

其中，编码端获取虚拟扬声器集合，编码端使用该虚拟扬声器集合对第一场景音频信号进行信号分解，以得到第一场景音频信号对应的第二主要声场成分。其中，第二主要声场成分表示的是第一场景音频信号中的主要声场所对应的音频信号。例如虚拟扬声器集合中包括多个虚拟扬声器，根据多个虚拟扬声器可以从第一场景音频信号中获取多个声场成分，即每个虚拟扬声器可以从第一场景音频信号中获取一个声场成分，接下来从多个声场成分中选择出第二主要声场成分，例如第二主要声场成分可以是多个声场成分中取值最大

的一个或几个声场成分，或第二主要声场成分可以是多个声场成分中方向占优的一个或几个声场成分。根据第二主要声场成分从虚拟扬声器集合中选择出第二目标虚拟扬声器，例如第二主要声场成分对应的虚拟扬声器就是编码端选择出的第二目标虚拟扬声器。本申请实施例中，编码端通过主要声场成分可以选择出第二目标虚拟扬声器，解决了编码端需要确定第二目标虚拟扬声器的问题。

其中，在本申请的一些实施例中，前述步骤 F1 根据第二主要声场成分从虚拟扬声器集合中选择出第二目标虚拟扬声器，包括：

根据第二主要声场成分从 HOA 系数集合中选择出与第二主要声场成分对应的 HOA 系数，HOA 系数集合中的 HOA 系数与虚拟扬声器集合中的虚拟扬声器一一对应；

确定虚拟扬声器集合中与第二主要声场成分对应的 HOA 系数对应的虚拟扬声器为第二目标虚拟扬声器。

其中，上述实现与前述实施例中确定第一目标虚拟扬声器的过程相类似，此处不再赘述。

其中，在本申请的一些实施例中，前述步骤 F1 根据第二主要声场成分从虚拟扬声器集合中选择出第二目标虚拟扬声器，还包括：

G1、根据第二主要声场成分获取第二目标虚拟扬声器的配置参数；

G2、根据第二目标虚拟扬声器的配置参数生成第二目标虚拟扬声器对应的 HOA 系数；

G3、确定虚拟扬声器集合中第二目标虚拟扬声器对应的 HOA 系数对应的虚拟扬声器为第二目标虚拟扬声器。

其中，上述实现与前述实施例中确定第一目标虚拟扬声器的过程相类似，此处不再赘述。

其中，上述实现与前述实施例中确定第一目标虚拟扬声器的过程相类似，此处不再赘述。

其中，在本申请的一些实施例中，步骤 G1 根据第二主要声场成分获取第二目标虚拟扬声器的配置参数，包括：

根据音频编码器的配置信息确定虚拟扬声器集合中的多个虚拟扬声器的配置参数；

根据第二主要声场成分从多个虚拟扬声器的配置参数中选择出第二目标虚拟扬声器的配置参数。

其中，上述实现与前述实施例中确定第一目标虚拟扬声器的配置参数的过程相类似，此处不再赘述。

其中，在本申请的一些实施例中，第二目标虚拟扬声器的配置参数包括：第二目标虚拟扬声器的位置信息和 HOA 阶数信息；

前述步骤 G2 根据第二目标虚拟扬声器的配置参数生成第二目标虚拟扬声器对应的 HOA 系数，包括：

根据第二目标虚拟扬声器的位置信息和 HOA 阶数信息确定第二目标虚拟扬声器对应的 HOA 系数。

其中，上述实现与前述实施例中确定第一目标虚拟扬声器对应的 HOA 系数的过程相类似，此处不再赘述。

在本申请的一些实施例中，第一场景音频信号，包括：待编码 HOA 信号；第二目标虚拟扬声器的属性信息包括第二目标虚拟扬声器的 HOA 系数；

步骤 D2 根据第一场景音频信号和第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号，包括：

对待编码 HOA 信号和第二目标虚拟扬声器的 HOA 系数进行线性组合，以得到第二虚拟扬声器信号。

在本申请的一些实施例中，第一场景音频信号包括：待编码高阶立体混响 HOA 信号；第二目标虚拟扬声器的属性信息包括第二目标虚拟扬声器的位置信息；

步骤 D2 根据第一场景音频信号和第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号，包括：

根据第二目标虚拟扬声器的位置信息获取第二目标虚拟扬声器对应的 HOA 系数；

对待编码 HOA 信号和第二目标虚拟扬声器对应的 HOA 系数进行线性组合，以得到第二虚拟扬声器信号。

其中，上述实现与前述实施例中确定第一虚拟扬声器信号的过程相类似，此处不再赘述。

在本申请实施例中，编码端生成第二虚拟扬声器信号之后，编码端还可以执行步骤 D3，对第二虚拟扬声器信号进行编码，并写入码流。其中，编码端所采用的编码方法与步骤 403 相类似，使得码流可以携带第二虚拟扬声器信号的编码结果。

其中，在本申请的一些实施例中，编码端执行的音频编码方法还可以包括如下步骤：

I1、对第一虚拟扬声器信号和第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号。

在执行步骤 I1 的场景下，相应地，步骤 D3 对第二虚拟扬声器信号进行编码包括：

对对齐后的第二虚拟扬声器信号进行编码；

相应地，步骤 403 对第一虚拟扬声器信号进行编码，包括：

对对齐后的第一虚拟扬声器信号进行编码。

其中，编码端可以生成第一虚拟扬声器信号和第二虚拟扬声器信号，编码端可以对第一虚拟扬声器信号和第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号，举例说明如下，有两个虚拟扬声器信号，当前帧的虚拟扬声器信号的声道顺序为 1、2，分别对应由目标虚拟扬声器 P1、P2 产生的虚拟扬声器信号，前一帧的虚拟扬声器信号的声道顺序为 1、2，分别对应由目标虚拟扬声器 P2、P1 产生的虚拟扬声器信号，则可以按照前一帧目标虚拟扬声器的顺序对当前帧虚拟扬声器信号的声道顺序进行调整，例如将当前帧的虚拟扬声器信号的声道顺序调整为 2、1，使得相同的目标虚拟扬声器产生的虚拟扬声器信号处于同一声道上。

编码端获取到对齐后的第一虚拟扬声器信号之后，可以对对齐后的第一虚拟扬声器信号进行编码，本申请实施例中通过将第一虚拟扬声器信号的各声道间重新调整对齐，增强了声道间相关性，有利于核心编码器对第一虚拟扬声器信号的编码处理。

在本申请的一些实施例中，编码端除了执行前述步骤之外，本申请实施例提供的音频编码方法还包括：

D1、根据第一场景音频信号从虚拟扬声器集合中选择出第二目标虚拟扬声器；

D2、根据第一场景音频信号和第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号。

相应地，在编码端执行步骤 D1 至 D2 的场景下，步骤 403 对第一虚拟扬声器信号进行编码，包括：

J1、根据第一虚拟扬声器信号和第二虚拟扬声器信号获得下混信号和边信息，边信息用于指示第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系；

J2、对下混信号以及边信息进行编码。

其中，编码端在获取到第一虚拟扬声器信号和第二虚拟扬声器信号之后，编码端还可以根据第一虚拟扬声器信号和第二虚拟扬声器信号进行下混处理，以生成下混信号，例如对第一虚拟扬声器信号和第二虚拟扬声器信号进行幅度上的下混处理，以得到下混信号。另外还可以根据第一虚拟扬声器信号和第二虚拟扬声器信号生成边信息，边信息用于指示第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系，该关系具有多种实现方式，该边信息可以用于解码端针对下混信号进行上混，以恢复出第一虚拟扬声器信号和第二虚拟扬声器信号。例如边信息包括信号信息丢失分析参数，以使得解码端通过信号信息丢失分析参数恢复出第一虚拟扬声器信号和第二虚拟扬声器信号。又如边信息具体可以是第一虚拟扬声器信号和第二虚拟扬声器信号的相关性参数，例如，可以是第一虚拟扬声器信号和第二虚拟扬声器信号的能量比例参数。以使得解码端通过上述相关性参数或者能量比例参数恢复出第一虚拟扬声器信号和第二虚拟扬声器信号。

其中，在本申请的一些实施例中，在编码端执行步骤 D1 至 D2 的场景下，编码端还可以执行如下步骤：

I1、对第一虚拟扬声器信号和第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号。

在执行步骤 I1 的场景下，相应地，步骤 J1 根据第一虚拟扬声器信号和第二虚拟扬声器信号获得下混信号和边信息，包括：

根据对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号获得下混信号和边信息；

相应的，边信息用于指示对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号之间的关系。

其中，编码端在生成下混信号之前，可以先执行虚拟扬声器信号的对齐操作，在完成对齐操作之后，再生成下混信号和边信息。本申请实施例中通过将第一虚拟扬声器信号和第二虚拟扬声器的各声道间重新调整对齐，增强了声道间相关性，有利于核心编码器对第一虚拟扬声器信号的编码处理。

需要说明的是，在本申请的上述实施例中，第二场景音频信号可以根据对齐前的第一虚拟扬声器信号和对齐前的第二虚拟扬声器信号获取，也可以根据对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号获取，具体实现方式取决于应用场景，此处不做限定。

在本申请的一些实施例中，在步骤 D1 根据第一场景音频信号从虚拟扬声器集合中选择

出第二目标虚拟扬声器前，本申请实施例提供的音频信号编码方法还包括：

K1、根据编码速率和/或第一场景音频信号的信号类型信息确定是否需要获取除第一目标虚拟扬声器以外的目标虚拟扬声器；

K2、若需要获取除第一目标虚拟扬声器以外的目标虚拟扬声器，才根据第一场景音频信号从虚拟扬声器集合中选择出第二目标虚拟扬声器。

其中，编码端还可以进行信号选择，以确定是否需要获取第二目标虚拟扬声器，在需要获取第二目标虚拟扬声器的情况下，编码端可以生成第二虚拟扬声器信号，在不需要获取第二目标虚拟扬声器的情况下，编码端可以不生成第二虚拟扬声器信号。其中，编码器可以根据音频编码器的配置信息和/或第一场景音频信号的信号类型信息进行决策，以确定在选择出第一目标虚拟扬声器之外是否还需要选择别的目标虚拟扬声器。例如，若编码速率高于预设的阈值，则确定需要获取两个主要声场成分对应的目标虚拟扬声器，则在确定出第一目标虚拟扬声器之外，还可以继续确定第二目标虚拟扬声器。又如，根据第一场景音频信号的信号类型信息确定需要获取包含声源方向占优的两个主要声场成分对应的目标虚拟扬声器，则在确定出第一目标虚拟扬声器之外，还可以继续确定第二目标虚拟扬声器。相反的，若根据编码速率和/或第一场景音频信号的信号类型信息确定只需要获取一个目标虚拟扬声器，则在确定第一目标虚拟扬声器之后，就确定不再获取除第一目标虚拟扬声器以外的目标虚拟扬声器。本申请实施例中通过信号选择，可以减少编码端进行编码的数据量，提高编码效率。

其中，编码端进行信号选择时，可以确定是否需要生成第二虚拟扬声器信号。由于编码端进行信号选择，会产生信息丢失，因此需要对不传输的虚拟扬声器信号进行信号补偿。信号补偿可以选择且不限于信息丢失分析，能量补偿，包络补偿，噪声补偿等。补偿的方法可以选择线性补偿或非线性补偿等。信号补偿之后可以生成边信息，该边信息可以被写入码流中，从而解码端可以通过码流得到边信息，解码端可以根据边信息进行信号补偿，从而提高解码端的解码信号质量。

通过前述实施例的举例说明，本申请实施例中可以根据第一场景音频信号和第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，音频编码端对该第一虚拟扬声器信号进行编码，而不再直接对第一场景音频信号进行编码，本申请实施例中根据第一场景音频信号选择出第一目标虚拟扬声器，基于该第一目标虚拟扬声器生成的第一虚拟扬声器信号可以表示空间中听音人所在的位置声场，该位置声场尽可能的接近录制第一场景音频信号时的原始声场，保证了音频编码端的编码质量，且对第一虚拟扬声器信号和残差信号进行编码以得到码流，该第一虚拟扬声器信号的编码数据量与第一目标虚拟扬声器有关，而与第一场景音频信号的声道个数无关，减少了编码数据量，提高编码效率。

在申请实施例中，编码端对虚拟扬声器信号进行编码，生成码流。然后编码端可以将该码流输出，并经过音频传输通道，发送至解码端。解码端执行后续步骤 411 至步骤 413。

411、接收码流。

其中，解码端从编码端接收码流。该码流可以携带编码后的第一虚拟扬声器信号。不限定的是，该码流还可以携带编码后的第一目标虚拟扬声器的属性信息。需要说明的是，码流中可以不携带第一目标虚拟扬声器的属性信息，此时解码端可以通过预先配置确定出

第一目标虚拟扬声器的属性信息。

另外，在本申请的一些实施例中，在编码端生成第二虚拟扬声器信号的情况下，该码流还可以携带第二虚拟扬声器信号。不限定的是，该码流还可以携带编码后的第二目标虚拟扬声器的属性信息。需要说明的是，码流中可以不携带第二目标虚拟扬声器的属性信息，此时解码端可以通过预先配置确定出第二目标虚拟扬声器的属性信息。

412、解码码流以获得虚拟扬声器信号。

其中，解码端接收到来自编码端的码流之后，对该码流进行解码，从该码流中得到虚拟扬声器信号。

需要说明的是，该虚拟扬声器信号具体可以是前述的第一虚拟扬声器信号，还可以是前述的第一虚拟扬声器信号和第二虚拟扬声器信号，此处不做限定。

在本申请的一些实施例中，解码端执行上述的步骤 411 至步骤 412 之后，本申请实施例提供的音频解码方法还包括如下步骤：

解码码流以获得目标虚拟扬声器的属性信息。

其中，编码端除了对虚拟扬声器进行编码，还可以对目标虚拟扬声器的属性信息进行编码，并将编码后的目标虚拟扬声器的属性信息写入到码流中，例如可以通过码流获取到第一目标虚拟扬声器的属性信息。本申请实施例中码流中可以携带编码后的第一目标虚拟扬声器的属性信息，使得解码端通过解码码流，就可以确定出第一目标虚拟扬声器的属性信息，便于解码端的音频解码。

413、根据目标虚拟扬声器的属性信息和虚拟扬声器信号获得重建的场景音频信号。

其中，解码端可以获取目标虚拟扬声器的属性信息，该目标虚拟扬声器为虚拟扬声器集合中用于回放重建的场景音频信号的虚拟扬声器。目标虚拟扬声器的属性信息可以包括目标虚拟扬声器的位置信息和目标虚拟扬声器的 HOA 系数。解码端获取到虚拟扬声器信号之后，解码端使用目标虚拟扬声器的属性信息进行信号重建，通过信号重建可以输出重建的场景音频信号。

在本申请的一些实施例中，目标虚拟扬声器的属性信息包括目标虚拟扬声器的 HOA 系数；

步骤 413 根据目标虚拟扬声器的属性信息以及虚拟扬声器信号获得重建的场景音频信号，包括：

对虚拟扬声器信号和目标虚拟扬声器的 HOA 系数进行合成处理，得到重建的场景音频信号。

其中，解码端首先确定目标虚拟扬声器的 HOA 系数，例如解码端中可以预先存储目标虚拟扬声器的HOA系数，解码端获取到虚拟扬声器信号和目标虚拟扬声器的HOA系数之后，根据虚拟扬声器信号和目标虚拟扬声器的 HOA 系数可以得到重建的场景音频信号。从而提高重建的场景音频信号的质量。

举例说明如下，目标虚拟扬声器的HOA系数用矩阵A’表示，矩阵A’的大小为$(M \times C)$，C为目标虚拟扬声器个数，M为N阶的HOA系数的声道个数。虚拟扬声器信号用矩阵W’表示，矩阵 W’ 的大小为$(C \times L)$，其中，L 为信号采样点个数。通过如下计算式得到重建的 HOA 信号：

$H = A'W'$，

通过上述计算式得到的 H 即为重建的 HOA 信号。

在本申请的一些实施例中，目标虚拟扬声器的属性信息包括目标虚拟扬声器的位置信息；

步骤 413 根据目标虚拟扬声器的属性信息以及虚拟扬声器信号获得重建的场景音频信号，包括：

根据目标虚拟扬声器的位置信息确定目标虚拟扬声器的 HOA 系数；

对虚拟扬声器信号和目标虚拟扬声器的 HOA 系数进行合成处理，得到重建的场景音频信号。

其中，目标虚拟扬声器的属性信息可以包括：目标虚拟扬声器的位置信息。解码端预先存储虚拟扬声器集合中每个虚拟扬声器的 HOA 系数，解码端还存储有每个虚拟扬声器的位置信息，例如解码端可以根据虚拟扬声器的位置信息和该虚拟扬声器的 HOA 系数之间的对应关系确定出目标虚拟扬声器的位置信息对应的 HOA 系数，或者解码端可以根据目标虚拟扬声器的位置信息计算出目标虚拟扬声器的 HOA 系数。因此解码端可以通过目标虚拟扬声器的位置信息确定目标虚拟扬声器的 HOA 系数。解决了解码端需要确定目标虚拟扬声器的 HOA 系数的问题。

在本申请的一些实施例中，通过编码端的方法说明可知，虚拟扬声器信号是根据第一虚拟扬声器信号和第二虚拟扬声器信号下混获得的下混信号。在这种实现场景下，本申请实施例提供的音频解码方法还包括：

解码码流以获得边信息，边信息用于指示第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系；

根据边信息和下混信号获得第一虚拟扬声器信号和第二虚拟扬声器信号。

其中，在本发明实施例中，所述第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系可以是直接关系，也可以是间接关系；例如在第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系为直接关系时，所述第一边信息可以包括第一虚拟扬声器信号和第二虚拟扬声器信号的相关性参数，例如可以是第一虚拟扬声器信号和第二虚拟扬声器信号的能量比例参数；例如在，第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系为间接关系时，所述第一边信息可以包括第一虚拟扬声器信号与下混信号之间的相关性参数，以及第二虚拟扬声器信号与下混信号之间的相关性参数，例如包括第一虚拟扬声器信号与下混信号之间的能量比例参数，以及第二虚拟扬声器信号与下混信号之间的能量比例参数。

在所述第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系可以是直接关系时，解码器可以根据下混信号，下混信号的获取方式以及该直接关系确定出第一虚拟扬声器信号和第二虚拟扬声器信号；在所述第一虚拟扬声器信号和第二虚拟扬声器信号之间的关系可以是间接关系时，解码器可以根据下混信号及该间接关系确定出第一虚拟扬声器信号和第二虚拟扬声器信号。

相应的，步骤 413 根据目标虚拟扬声器的属性信息以及虚拟扬声器信号获得重建的场景音频信号，包括：

根据目标虚拟扬声器的属性信息、第一虚拟扬声器信号和第二虚拟扬声器信号获得重

建的场景音频信号。

其中，编码端根据第一虚拟扬声器信号和第二虚拟扬声器信号进行下混处理时生成下混信号，编码端还可以针对下混信号进行信号补偿，以生成边信息，该边信息可以被写入码流中，解码端可以通过码流得到边信息，解码端可以根据边信息进行信号补偿，以得到第一虚拟扬声器信号和第二虚拟扬声器信号，因此在进行信号重建时，可以使用第一虚拟扬声器信号和第二虚拟扬声器信号，以及前述的目标虚拟扬声器的属性信息，从而提高解码端的解码信号质量。

通过前述实施例的举例说明，本申请实施例中可以从码流中解码得到虚拟扬声器信号，虚拟扬声器信号作为场景音频信号的回放信号，通过目标虚拟扬声器的属性信息和虚拟扬声器信号得到了重建的场景音频信号，本申请实施例中，获取到的码流中携带虚拟扬声器信号和残差信号，减少了解码的数据量，提高了解码效率。

举例说明如下，本申请实施例中，第一虚拟扬声器信号相比于第一场景音频信号，采用较少的声道进行表示，例如第一场景音频信号为3阶HOA信号，该HOA信号为16个声道，本申请实施例中可以将16个声道压缩为2个声道，即编码端生成的虚拟扬声器信号为2个声道，例如编码端生成的虚拟扬声器信号可以包括前述的第一虚拟扬声器信号和第二虚拟扬声器信号等，编码端生成的虚拟扬声器信号的声道数与第一场景音频信号的声道数无关。通过后续步骤描述可知，码流中可以携带2个声道的虚拟扬声器信号，相应的，解码端接收该码流，解码该码流得到的虚拟扬声器信号为2个声道，解码端通过2个声道的虚拟扬声器信号可以重建出16个声道的场景音频信号，且保证了重建的场景音频信号与原始的场景音频信号相比时，具有主观和客观质量相当的效果。

为便于更好的理解和实施本申请实施例的上述方案，下面举例相应的应用场景来进行具体说明。

本申请实施例中以场景音频信号为 HOA 信号为例，声波在理想介质中传播，波数为 $k = w/c$，角频率 $w = 2\pi f$，f 为声波频率，c 为声速。则声压 p满足如下计算式，其中∇^2为拉普拉斯算子：

$$\nabla^2 p + k^2 p = 0。$$

在球坐标下求上述等式方程，在无源球形区域内，该方程解为如下计算式：

$$p(r，\ \theta，\ \varphi，\ k) = s\sum_{m=0}^{\infty}(2m+1)j^m j_m^{kr}(kr)\sum_{0\le n\le m，\ \sigma=\pm 1} Y_{m，\ n}^{\sigma}(\theta_s，\ \varphi_s)Y_{m，\ n}^{\sigma}(\theta，\ \varphi)。$$

在上述计算式中，r 表示球半径，θ表示水平角，φ表示仰角，k表示波数，s为理想平面波的幅度，m为 HOA 阶数序号，$j^m j_m^{kr}(kr)$是球贝塞尔函数，又称径向基函数，其中第一个j是虚数单位。$(2m+1)j^m j_m^{kr}(kr)$不随角度变化。$Y_{m，\ n}^{\sigma}(\theta，\ \varphi)$即为$\theta$，$\varphi$方向的球谐函数，$Y_{m，\ n}^{\sigma}(\theta_s，\ \varphi_s)$是声源方向的球谐函数。

HOA 系数可以表述为：$B_{m，\ n}^{\sigma} = s \cdot Y_{m，\ n}^{\sigma}(\theta_s，\ \varphi_s)$。

进而给出如下计算式：

$p(r，\theta，\varphi，k)=\sum_{m=0}^{\infty}j^{m}j_{m}^{kr}(kr)\sum_{0\le n\le m，\sigma=\pm1}B_{m，n}^{\sigma}Y_{m，n}^{\sigma}(\theta，\varphi)$。

上述计算式表明声场可以在球面上按球谐函数展开，使用系数 $B_{m，n}^{\sigma}$ 进行表示。或者，已知系数 $B_{m，n}^{\sigma}$ 就可以重建声场。将上式截断到第 N 项，以系数 $B_{m，n}^{\sigma}$ 作为对声场的近似描述，则称为 N 阶的 HOA 系数，该 HOA 系数也可以称为 Ambisonic 系数。 N 阶的 HOA 系数共有 $(N+1)^2$个声道。其中，一阶以上的 Ambisonic 信号也称为 HOA 信号。将球谐函数按照 HOA 信号一个采样点对应的系数进行叠加，就能实现该采样点对应的时刻空间声场的重构。

例如，在一种配置下，HOA 阶数可以为 2 至 6 阶，对场景音频录制时信号采样率为 48 至 192 kHz， 采样深度为 16 或 24Bit。HOA 信号的特点是带有声场的空间信息，是空间某点声场信号一定精度的描述。因此，可以考虑使用另一种表示形式描述该点的声场信号，如果这种描述方法能够使用更少的数据量对该点信号达到同样精确度的描述，就能达到信号压缩的目的。

空间声场可以分解为多个平面波的叠加。因此，可以将 HOA 信号表达的声场，重新使用多个平面波的叠加来表达，每个平面波使用一个声道的音频信号和一个方向向量表示。如果平面波叠加的表示形式能够使用更少的声道数目较好的表达原始声场，则可以达到信号压缩的目的。

HOA 信号在实际回放时，可以通过耳机回放，也可以通过布置在房间中的多个扬声器回放。使用扬声器回放时，基本方法是通过多个扬声器的声场的叠加，使得空间中某点（听音人所在的位置）声场在某个标准下尽可能的接近录制 HOA 信号时的原始声场。本申请实施例假设一个虚拟扬声器阵列，然后计算该虚拟扬声器阵列的回放信号，使用该回放信号作为传输信号，并进而生成压缩后的信号。解码端通过对码流进行解码，得到该回放信号，并通过该回放信号重建出场景音频信号。

本申请实施例提供适用于场景音频信号编码的编码端，和适用于场景音频信号解码的解码端。其中，编码端将原始 HOA 信号编码为压缩码流，编码端向解码端发送该压缩码流，然后解码端将压缩码流恢复为重建 HOA 信号。本申请实施例中，编码端进行压缩后的数据量尽可能小，或在同等码率下解码端重建后得到的 HOA 信号的质量更高。

本申请实施例可以解决编码 HOA 信号的时候，数据量大，带宽占用高，压缩效率较低，编码质量不高的问题。由于 N 阶的 HOA 信号具有 $(N+1)^2$个声道，直接传输该 HOA 信号需要消耗较大的带宽，因此需要一种有效的多声道编码方案。

本申请实施例采取了不同的声道提取方法，且本申请实施例中对声源的假设不做限定，不依赖时频域点单声源假设，可以更有效的处理多声源信号等复杂场景。本申请实施例的编解码器提供一种采用较少的声道用于表示原始HOA信号的空间编解码方法。如图 5 所示，为本申请实施例提供的编码端的一种结构示意图，编码端包括空间编码器和核心编码器，其中，空间编码器可以对待编码 HOA 信号进行声道提取，以生成虚拟扬声器信号，核心编码器可以对虚拟扬声器信号进行编码，以得到码流，编码端向解码端发送码流。如图 6 所示，为本申请实施例提供的解码端的一种结构示意图，解码端包括：核心解码器和空间解码器，其中，核心解码器先接收到来自编码端的码流，然后从该码流中解码出虚拟扬声器信号，接下来空间解码器对该虚拟扬声器信号进行重建，以得到重建的 HOA 信号。

接下来分别从编码端和解码端进行举例说明。

如图 7 所示，首先对本申请实施例提供的编码端进行说明，该编码端可以包括：虚拟扬声器配置单元、编码分析单元、虚拟扬声器集合生成单元、虚拟扬声器选择单元、虚拟扬声器信号生成单元和核心编码器处理单元。接下来分别对编码端的各个组成单元的功能进行说明。本申请实施例中，图 7 所示的编码端可以生成一个虚拟扬声器信号，也可以生成多个虚拟扬声器信号，其中，多个虚拟扬声器信号的生成流程可以是根据图 7 所示的编码器结构进行多次生成，接下来以一个虚拟扬声器信号的生成流程为例。

虚拟扬声器配置单元，用于对虚拟扬声器集合中的虚拟扬声器进行配置，以得到多个虚拟扬声器。

虚拟扬声器配置单元根据编码器配置信息输出虚拟扬声器配置参数。编码器配置信息包括且不限于：HOA 阶数，编码比特率，用户自定义信息等，虚拟扬声器配置参数包括且不限于：虚拟扬声器的个数，虚拟扬声器的 HOA 阶数、虚拟扬声器的位置坐标等。

虚拟扬声器配置单元输出的虚拟扬声器配置参数作为虚拟扬声器集合生成单元的输入。

编码分析单元，用于对待编码 HOA 信号进行编码分析，例如分析待编码 HOA 信号的声场分布，包括待编码 HOA 信号的声源个数、方向性、弥散度等特征，作为决定如何选择目标虚拟扬声器的判断条件之一。

不限定的是，本申请实施例中，编码端中还可以不包括编码分析单元，即编码端可以不对输入信号进行分析，则采用一种默认配置决定如何选择目标虚拟扬声器。

其中，编码端获取待编码 HOA 信号，例如可以将从实际采集设备记录的 HOA 信号或采用人工音频对象合成的 HOA 信号作为编码器的输入，同时编码器输入的待编码 HOA 信号可以是时域 HOA 信号也可以是频域 HOA 信号。

虚拟扬声器集合生成单元，用于生成虚拟扬声器集合，该虚拟扬声器集合中可以包括：多个虚拟扬声器，虚拟扬声器集合中的虚拟扬声器也可以称为“候选虚拟扬声器”。

虚拟扬声器集合生成单元生成指定的候选虚拟扬声器 HOA 系数。生成候选虚拟扬声器 HOA 系数需要候选虚拟扬声器的坐标（即位置坐标或者位置信息）和候选虚拟扬声器的 HOA 阶数，候选虚拟扬声器的坐标确定方法包括且不限于按等距规则产生 K 个虚拟扬声器、根据听觉感知原理生成非均匀分布的 K 个候选虚拟扬声器，以下举例一种产生均匀分布固定个数虚拟扬声器的方法。

根据候选虚拟扬声器的个数生成分布均匀的候选虚拟扬声器的坐标，例如使用数值迭代计算方法给出近似均匀的扬声器排布。如图 8 所示，为球面上近似均匀分布的虚拟扬声器的示意图，设想在单位球面上分布一些质点，并在这些质点之间设置二次方反比的斥力，与同种电荷之间的静电斥力类似。让这些质点在斥力作用下自由运动，可以期望在其达到稳态时，质点的分布应该趋于均匀。计算中，对实际的物理规律进行简化，直接令质点的移动距离等于受力。则对于第 i个质点，其在迭代计算的某一步的运动距离也即受到的虚拟力为如下计算式：

$$\vec{D} = \vec{F} = \sum_{j=1,\ j\neq i}^{N} \frac{k}{r_{ij}^2}\vec{d}_{ij}。$$

其中，$\vec{D}$ 代表位移矢量，$\vec{F}$ 代表力矢量，r_{ij} 代表第 i个质点与第 j个质点之间的距离，

$\vec{d}_{ij}$ 代表从第 j个质点指向第i个质点的方向矢量。参数 k 控制单步步长的大小，质点的初始位置随机指定即可。

质点按照位移矢量 $\vec{D}$ 运动后，一般会偏离单位球面。在下一步迭代前，通过归一化质点与球心的距离，将其移动回单位球面即可，由此可以得到如图 8 所示的虚拟扬声器分布示意图，球面上近似均匀分布有多个虚拟扬声器。

接下来生成候选虚拟扬声器 HOA 系数。振幅为 s，扬声器位置坐标为(θ_s, φ_s)的理想平面波，使用球谐函数展开后的形式为如下计算式：

$$p(r, \theta, \varphi, k) =$$

$$s\sum_{m=0}^{\infty}(2m+1)j^m j_m^{kr}(kr)\sum_{0\le n\le m,\ \sigma=\pm 1} Y_{m,n}^{\sigma}(\theta_s, \varphi_s)Y_{m,n}^{\sigma}(\theta, \varphi)。$$

对于平面波的 HOA 系数为$B_{m,n}^{\sigma}$，满足如下计算式：

$$B_{m,n}^{\sigma} = s \cdot Y_{m,n}^{\sigma}(\theta_s, \varphi_s)。$$

虚拟扬声器集合生成单元输出的候选虚拟扬声器的 HOA 系数作为虚拟扬声器选择单元的输入。

虚拟扬声器选择单元，用于根据待编码 HOA 信号从虚拟扬声器集合中的多个候选虚拟扬声器中选择出目标虚拟扬声器，该目标虚拟扬声器可以称为“与待编码 HOA 信号匹配的虚拟扬声器”，或者简称为匹配虚拟扬声器。

虚拟扬声器选择单元将待编码 HOA 信号与虚拟扬声器集合生成单元输出的候选虚拟扬声器 HOA 系数匹配，选择出指定的匹配虚拟扬声器。

接下来对虚拟扬声器的选择方法进行举例说明，一种实施例中，得到候选虚拟扬声器后，将待编码 HOA 信号与虚拟扬声器集合生成单元输出的候选虚拟扬声器 HOA 系数进行匹配，寻找待编码 HOA 信号在候选虚拟扬声器上的最佳匹配，目标是使用候选虚拟扬声器 HOA 系数匹配组合待编码 HOA 信号。一种实施例中，使用候选虚拟扬声器 HOA 系数与待编码 HOA 信号做内积，选取内积绝对值最大的候选虚拟扬声器为目标虚拟扬声器，即匹配虚拟扬声器，并将待编码 HOA 信号在该候选虚拟扬声器的投影叠加到该候选虚拟扬声器 HOA 系数的线性组合上，然后将投影向量从待编码 HOA 信号中减去得到差值，对差值重复上述过程实现迭代计算，每迭代一次产生一个匹配虚拟扬声器，输出匹配虚拟扬声器坐标和匹配虚拟扬声器 HOA 系数。可以理解的是，匹配虚拟扬声器会选取多个，每迭代一次产生一个匹配虚拟扬声器。

虚拟扬声器选择单元输出的目标虚拟扬声器的坐标和目标虚拟扬声器的 HOA 系数作为虚拟扬声器信号生成单元的输入。

在本申请的一些实施例中，编码端除了包括图 7 所示的组成单元之外，还可以包括边信息生成单元。不限定的是，编码端还可以不包括边信息生成单元，此处仅为举例。

虚拟扬声器选择单元输出的目标虚拟扬声器的坐标和/或目标虚拟扬声器的 HOA 系数作为边信息生成单元的输入。

边信息生成单元将目标虚拟扬声器的 HOA 系数或目标虚拟扬声器的坐标转换为边信息，

利于核心编码器的处理和传输。

边信息生成单元的输出作为核心编码器处理单元的输入。

虚拟扬声器信号生成单元，用于根据待编码 HOA 信号和目标虚拟扬声器的属性信息生成虚拟扬声器信号。

虚拟扬声器信号生成单元通过待编码 HOA 信号和目标虚拟扬声器的 HOA 系数计算虚拟扬声器信号。

匹配虚拟扬声器 HOA 系数用矩阵 A 表示，用矩阵 A 可以线性组合出待编码 HOA 信号，其中可以采用最小二乘方法求得理论的最优解w，即为虚拟扬声器信号，例如可以采用如下计算式：

$w = A^{-1}X$，

其中，A^{-1} 代表矩阵 A 的逆矩阵，矩阵 A 的大小为$(M \times C)$，C 为目标虚拟扬声器个数，M 为N阶的 HOA 系数的声道个数，a 表示目标虚拟扬声器的 HOA 系数，例如，

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1C} \\ \vdots & & \vdots \\ a_{M1} & \cdots & a_{MC} \end{bmatrix}。$$

其中，X 代表待编码 HOA 信号，矩阵 X 的大小为(M×L)，M 为N阶的 HOA 系数的声道个数，L 为采样点数，x 表示待编码 HOA 信号的系数，例如，

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1L} \\ \vdots & & \vdots \\ x_{M1} & \cdots & x_{ML} \end{bmatrix}。$$

虚拟扬声器信号生成单元输出的虚拟扬声器信号作为核心编码器处理单元的输入。

在本申请的一些实施例中，编码端除了包括图 7 所示的组成单元之外，还可以包括信号对齐单元。不限定的是，编码端还可以不包括信号对齐单元，此处仅为举例。

虚拟扬声器信号生成单元输出的虚拟扬声器信号作为信号对齐单元的输入。

信号对齐单元，用于将虚拟扬声器信号各声道间重新调整，增强声道间相关性，利于核心编码器处理。

信号对齐单元输出的对齐后的虚拟扬声器信号为核心编码器处理单元的输入。

核心编码器处理单元，用于对边信息和对齐后的虚拟扬声器信号进行核心编码器处理，得到传输码流。

核心编码器处理包括且不限于变换、量化、心理声学模型、码流产生等，可以对频域声道进行处理也可以对时域声道进行处理，此处不做限定。

如图 9 所示，本申请实施例提供的解码端可包含：核心解码器处理单元和 HOA 信号重建单元。

核心解码器处理单元，用于对传输码流进行核心解码器处理，得到虚拟扬声器信号。

不限的是，若编码端在码流中携带边信息，则解码端还需要包括：边信息解码单元。

边信息解码单元，用于对核心解码器处理单元输出的解码边信息进行解码，以得到解码后的边信息。

核心解码器处理可以包括变换、码流解析、反量化等，可以对频域声道进行处理也可以对时域声道进行处理，此处不做限定。

核心解码器处理单元输出的虚拟扬声器信号为 HOA 信号重建单元的输入，核心解码器处理单元输出的解码边信息为边信息解码单元的输入。

边信息解码单元将解码边信息转为目标虚拟扬声器的 HOA 系数。

边信息解码单元输出的目标虚拟扬声器的 HOA 系数为 HOA 信号重建单元的输入。

HOA 信号重建单元，用于通过虚拟扬声器信号和目标虚拟扬声器的 HOA 系数对 HOA 信号进行重建。

目标虚拟扬声器的 HOA 系数用于矩阵A'表示，矩阵 A'的大小为$(M \times C)$，记为 A'，C 为目标虚拟扬声器个数，M 为N 阶的 HOA 系数的声道个数。虚拟扬声器信号构成$(C \times L)$ 矩阵，记为 W'，其中 L 为信号采样点个数，通过如下计算式得到重建的 HOA 信号 H：

$H = A'$ W'，

其中，HOA 信号重建单元输出的重建的 HOA 信号为解码端的输出。

本申请实施例中，编码端可以利用空间编码器，将原始 HOA 信号采用较少的声道进行表示，例如原始 3 阶 HOA 信号，采用本申请实施例的空间编码器可以将 16 个声道压缩为 4 个声道，且保证了主观听力无明显差别。其中，主观听力测试是音频编解码中的一种评价标准，无明显差别是主观评价的一种等级。

在本申请的另一些实施例中，编码端的虚拟扬声器选择单元从虚拟扬声器集合中选择出目标虚拟扬声器，还可以采用指定方位的虚拟扬声器作为目标虚拟扬声器，虚拟扬声器信号生成单元直接在各个目标虚拟扬声器上做投影得到虚拟扬声器信号。

在上述方式中，通过指定方位的虚拟扬声器作为目标虚拟扬声器，可以简化虚拟扬声器选择过程，可以提高编解码速度。

在本申请的另一些实施例中，编码器端可以不包括信号对齐单元，此时虚拟扬声器信号生成单元的输出直接进行核心编码器进行编码处理。通过上述方式，减少信号对齐处理，降低编码器端复杂度。

通过前述的举例说明可知，本申请实施例将选择出的目标虚拟扬声器应用于HOA信号编解码上，本申请实施例能够得到准确的HOA信号声源定位，重建HOA信号方向更为准确，编码效率更高，且解码端复杂度非常低，利于移动端应用，可以提升编解码的性能。

需要说明的是，对于前述的各方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请并不受所描述的动作顺序的限制，因为依据本申请，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定是本申请所必须的。

为便于更好的实施本申请实施例的上述方案，下面还提供用于实施上述方案的相关装置。

请参阅图 10 所示，本申请实施例提供的一种音频编码装置 1000，可以包括：获取模块 1001、信号生成模块 1002 和编码模块 1003，其中，

获取模块，用于根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器；

信号生成模块，用于根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号；

编码模块，用于对所述第一虚拟扬声器信号进行编码，以得到码流。

在本申请的一些实施例中，所述获取模块，用于根据所述虚拟扬声器集合从所述当前场景音频信号中获取主要声场成分；根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器。

在本申请的一些实施例中，所述获取模块，用于根据所述主要声场成分从高阶立体混响 HOA 系数集合中选择出与所述主要声场成分对应的 HOA 系数，所述 HOA 系数集合中的 HOA 系数与所述虚拟扬声器集合中的虚拟扬声器一一对应；确定所述虚拟扬声器集合中与所述主要声场成分对应的 HOA 系数对应的虚拟扬声器为所述第一目标虚拟扬声器。

在本申请的一些实施例中，所述获取模块，用于根据所述主要声场成分获取所述第一目标虚拟扬声器的配置参数；根据所述第一目标虚拟扬声器的配置参数生成所述第一目标虚拟扬声器对应的 HOA 系数；确定所述虚拟扬声器集合中所述第一目标虚拟扬声器对应的 HOA 系数对应的虚拟扬声器为所述目标虚拟扬声器。

在本申请的一些实施例中，所述获取模块，用于根据音频编码器的配置信息确定所述虚拟扬声器集合中的多个虚拟扬声器的配置参数；根据所述主要声场成分从所述多个虚拟扬声器的配置参数中选择出所述第一目标虚拟扬声器的配置参数。

在本申请的一些实施例中，所述第一目标虚拟扬声器的配置参数包括：所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息；

所述获取模块，用于根据所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息确定所述第一目标虚拟扬声器对应的 HOA 系数。

在本申请的一些实施例中，所述编码模块，还用于对所述第一目标虚拟扬声器的属性信息进行编码，并写入所述码流。

在本申请的一些实施例中，所述当前场景音频信号，包括：待编码 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的 HOA 系数；

所述信号生成模块，用于对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

在本申请的一些实施例中，所述当前场景音频信号包括：待编码高阶立体混响 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的位置信息；

所述信号生成模块，用于根据所述第一目标虚拟扬声器的位置信息获取所述第一目标虚拟扬声器对应的 HOA 系数；对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

在本申请的一些实施例中，所述获取模块，用于根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

所述信号生成模块，用于根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

所述编码模块，用于对所述第二虚拟扬声器信号进行编码，并写入所述码流。

在本申请的一些实施例中，所述信号生成模块，用于对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应地，所述编码模块，用于对所述对齐后的第二虚拟扬声器信号进行编码；

相应地，所述编码模块，用于对所述对齐后的第一虚拟扬声器信号进行编码。

在本申请的一些实施例中，所述获取模块，用于根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

所述信号生成模块，用于根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

相应地，所述编码模块，用于根据所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得下混信号和边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；对所述下混信号以及所述边信息进行编码。

在本申请的一些实施例中，所述信号生成模块，用于对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应的，所述编码模块，用于根据所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号获得所述下混信号和所述边信息；

相应的，所述边信息用于指示所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号之间的关系。

在本申请的一些实施例中，所述获取模块，用于在根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器前，根据编码速率和/或所述当前场景音频信号的信号类型信息确定是否需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器；若需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器，才根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器。

请参阅图 11 所示，本申请实施例提供的一种音频解码装置 1100，可以包括：接收模块 1101、解码模块 1102、重建模块 1103，其中，

接收模块，用于接收码流；

解码模块，用于解码所述码流以获得虚拟扬声器信号；

重建模块，用于根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号。

在本申请的一些实施例中，所述解码模块，还用于解码所述码流以获得所述目标虚拟扬声器的属性信息。

在本申请的一些实施例中，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的高阶立体混响 HOA 系数；

所述重建模块，用于对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合

成处理，以获得所述重建的场景音频信号。

在本申请的一些实施例中，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的位置信息；

所述重建模块，用于根据所述目标虚拟扬声器的位置信息确定所述目标虚拟扬声器的HOA 系数；对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

在本申请的一些实施例中，所述虚拟扬声器信号是根据第一虚拟扬声器信号和第二虚拟扬声器信号下混获得的下混信号，所述装置还包括：信号补偿模块，其中，

所述解码模块，用于解码所述码流以获得边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；

所述信号补偿模块，用于根据所述边信息和所述下混信号获得所述第一虚拟扬声器信号和所述第二虚拟扬声器信号；

相应的，所述重建模块，用于根据所述目标虚拟扬声器的属性信息、所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得所述重建的场景音频信号

需要说明的是，上述装置各模块/单元之间的信息交互、执行过程等内容，由于与本申请方法实施例基于同一构思，其带来的技术效果与本申请方法实施例相同，具体内容可参见本申请前述所示的方法实施例中的叙述，此处不再赘述。

本申请实施例还提供一种计算机存储介质，其中，该计算机存储介质存储有程序，该程序执行包括上述方法实施例中记载的部分或全部步骤。

接下来介绍本申请实施例提供的另一种音频编码装置，请参阅图 12 所示，音频编码装置 1200 包括：

接收器 1201、发射器 1202、处理器 1203 和存储器 1204（其中音频编码装置 1200 中的处理器 1203 的数量可以一个或多个，图 12 中以一个处理器为例）。在本申请的一些实施例中，接收器 1201、发射器 1202、处理器 1203 和存储器 1204 可通过总线或其它方式连接，其中，图 12 中以通过总线连接为例。

存储器 1204 可以包括只读存储器和随机存取存储器，并向处理器 1203 提供指令和数据。存储器 1204 的一部分还可以包括非易失性随机存取存储器（non-volatile random access memory，NVRAM）。存储器 1204 存储有操作系统和操作指令、可执行模块或者数据结构，或者它们的子集，或者它们的扩展集，其中，操作指令可包括各种操作指令，用于实现各种操作。操作系统可包括各种系统程序，用于实现各种基础业务以及处理基于硬件的任务。

处理器 1203 控制音频编码装置的操作，处理器 1203 还可以称为中央处理单元（central processing unit，CPU）。具体的应用中，音频编码装置的各个组件通过总线系统耦合在一起，其中总线系统除包括数据总线之外，还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见，在图中将各种总线都称为总线系统。

上述本申请实施例揭示的方法可以应用于处理器 1203 中，或者由处理器 1203 实现。处理器 1203 可以是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器 1203 中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处

理器 1203 可以是通用处理器、数字信号处理器（digital signal processing，DSP）、专用集成电路（application specific integrated circuit，ASIC）、现场可编程门阵列（field-programmable gate array，FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本申请实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器 1204，处理器 1203 读取存储器 1204 中的信息，结合其硬件完成上述方法的步骤。

接收器 1201 可用于接收输入的数字或字符信息，以及产生与音频编码装置的相关设置以及功能控制有关的信号输入，发射器 1202 可包括显示屏等显示设备，发射器 1202 可用于通过外接接口输出数字或字符信息。

本申请实施例中，处理器 1203 用于执行前述实施例图 4 所示的由音频编码装置执行的音频编码方法。

接下来介绍本申请实施例提供的另一种音频解码装置，请参阅图 13 所示，音频解码装置 1300 包括：

接收器 1301、发射器 1302、处理器 1303 和存储器 1304（其中音频解码装置 1300 中的处理器 1303 的数量可以一个或多个，图 13 中以一个处理器为例）。在本申请的一些实施例中，接收器 1301、发射器 1302、处理器 1303 和存储器 1304 可通过总线或其它方式连接，其中，图 13 中以通过总线连接为例。

存储器 1304 可以包括只读存储器和随机存取存储器，并向处理器 1303 提供指令和数据。存储器 1304 的一部分还可以包括 NVRAM。存储器 1304 存储有操作系统和操作指令、可执行模块或者数据结构，或者它们的子集，或者它们的扩展集，其中，操作指令可包括各种操作指令，用于实现各种操作。操作系统可包括各种系统程序，用于实现各种基础业务以及处理基于硬件的任务。

处理器 1303 控制音频解码装置的操作，处理器 1303 还可以称为 CPU。具体的应用中，音频解码装置的各个组件通过总线系统耦合在一起，其中总线系统除包括数据总线之外，还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见，在图中将各种总线都称为总线系统。

上述本申请实施例揭示的方法可以应用于处理器 1303 中，或者由处理器 1303 实现。处理器 1303 可以是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器 1303 中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器 1303 可以是通用处理器、DSP、ASIC、FPGA 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本申请实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读

存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器 1304，处理器 1303 读取存储器 1304 中的信息，结合其硬件完成上述方法的步骤。

本申请实施例中，处理器 1303，用于执行前述实施例图 4 所示的由音频解码装置执行的音频解码方法。

在另一种可能的设计中，当音频编码装置或者音频解码装置为终端内的芯片时，芯片包括：处理单元和通信单元，所述处理单元例如可以是处理器，所述通信单元例如可以是输入/输出接口、管脚或电路等。该处理单元可执行存储单元存储的计算机执行指令，以使该终端内的芯片执行上述第一方面任意一项的音频编码方法，或者第二方面任意一项的音频解码方法。可选地，所述存储单元为所述芯片内的存储单元，如寄存器、缓存等，所述存储单元还可以是所述终端内的位于所述芯片外部的存储单元，如只读存储器（read-onlymemory，ROM)或可存储静态信息和指令的其他类型的静态存储设备，随机存取存储器（randomaccessmemory，RAM)等。

其中，上述任一处提到的处理器，可以是一个通用中央处理器，微处理器，ASIC，或一个或多个用于控制上述第一方面或第二方面方法的程序执行的集成电路。

另外需说明的是，以上所描述的装置实施例仅仅是示意性的，其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。另外，本申请提供的装置实施例附图中，模块之间的连接关系表示它们之间具有通信连接，具体可以实现为一条或多条通信总线或信号线。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件的方式来实现，当然也可以通过专用硬件包括专用集成电路、专用 CPU、专用存储器、专用元器件等来实现。一般情况下，凡由计算机程序完成的功能都可以很容易地用相应的硬件来实现，而且，用来实现同一功能的具体硬件结构也可以是多种多样的，例如模拟电路、数字电路或专用电路等。但是，对本申请而言更多情况下软件程序实现是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在可读取的存储介质中，如计算机的软盘、U 盘、移动硬盘、ROM、RAM、磁碟或者光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施例所述的方法。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。

所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例

如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存储的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘（Solid State Disk，SSD））等。

权 利 要 求

1.一种音频编码方法，其特征在于，包括：

根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器；

根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号；

对所述第一虚拟扬声器信号进行编码，以得到码流。

2.根据权利要求 1 所述的方法，其特征在于，所述方法还包括：

根据所述虚拟扬声器集合从所述当前场景音频信号中获取主要声场成分；

所述根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器包括：

根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器。

3.根据权利要求 2 所述的方法，其特征在于，所述根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器，包括：

根据所述主要声场成分从高阶立体混响 HOA 系数集合中选择出与所述主要声场成分对应的 HOA 系数，所述 HOA 系数集合中的 HOA 系数与所述虚拟扬声器集合中的虚拟扬声器一一对应；

确定所述虚拟扬声器集合中与所述主要声场成分对应的 HOA 系数对应的虚拟扬声器为所述第一目标虚拟扬声器。

4.根据权利要求 2 所述的方法，其特征在于，所述根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器，包括：

根据所述主要声场成分获取所述第一目标虚拟扬声器的配置参数；

根据所述第一目标虚拟扬声器的配置参数生成所述第一目标虚拟扬声器对应的 HOA 系数；

确定所述虚拟扬声器集合中所述第一目标虚拟扬声器对应的 HOA 系数对应的虚拟扬声器为所述目标虚拟扬声器。

5.根据权利要求 4 所述的方法，其特征在于，所述根据所述主要声场成分获取所述第一目标虚拟扬声器的配置参数，包括：

根据音频编码器的配置信息确定所述虚拟扬声器集合中的多个虚拟扬声器的配置参数；

根据所述主要声场成分从所述多个虚拟扬声器的配置参数中选择出所述第一目标虚拟扬声器的配置参数。

6.根据权利要求 4 或 5 所述的方法，其特征在于，所述第一目标虚拟扬声器的配置参数包括：所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息；

所述根据所述第一目标虚拟扬声器的配置参数生成所述第一目标虚拟扬声器对应的 HOA 系数，包括：

根据所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息确定所述第一目标虚拟扬声器对应的 HOA 系数。

7.根据权利要求 1 至 6 中任一项所述的方法，其特征在于，所述方法还包括：

对所述第一目标虚拟扬声器的属性信息进行编码，并写入所述码流。

8.根据权利要求 1 至 7 中任一项所述的方法，其特征在于，所述当前场景音频信号包括：待编码高阶立体混响 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的 HOA 系数；

所述根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，包括：

对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

9.根据权利要求 1 至 7 任一项所述的方法，其特征在于，所述当前场景音频信号包括：待编码高阶立体混响 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的位置信息；

所述根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号，包括：

根据所述第一目标虚拟扬声器的位置信息获取所述第一目标虚拟扬声器对应的 HOA 系数；

对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

10.根据权利要求 1 至 9 中任一项所述的方法，其特征在于，所述方法还包括：

根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

对所述第二虚拟扬声器信号进行编码，并写入所述码流。

11.根据权利要求 10 所述的方法，其特征在于，所述方法还包括：

对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应地，所述对所述第二虚拟扬声器信号进行编码包括：

对所述对齐后的第二虚拟扬声器信号进行编码；

相应地，所述对所述第一虚拟扬声器信号进行编码，包括：

对所述对齐后的第一虚拟扬声器信号进行编码。

12.根据权利要求 1 至 9 中任一项所述的方法，其特征在于，所述方法还包括：

根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

相应地，所述对所述第一虚拟扬声器信号进行编码，包括：

根据所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得下混信号和边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；

对所述下混信号以及所述边信息进行编码。

13.根据权利要求 12 所述的方法，其特征在于，所述方法还包括：

对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后

的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应的，所述根据所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得下混信号和边信息，包括：

根据所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号获得所述下混信号和所述边信息；

相应的，所述边信息用于指示所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号之间的关系。

14. 根据权利要求 10 至 13 任一项所述的方法，其特征在于，在根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器前，所述方法还包括：

根据编码速率和/或所述当前场景音频信号的信号类型信息确定是否需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器；

若需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器，才根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器。

15. 一种音频解码方法，其特征在于，包括：

接收码流；

解码所述码流以获得虚拟扬声器信号；

根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号。

16. 根据权利要求 15 所述的方法，其特征在于，所述方法还包括：

解码所述码流以获得所述目标虚拟扬声器的属性信息。

17. 根据权利要求 16 所述的方法，其特征在于，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的高阶立体混响 HOA 系数；

所述根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号，包括：

对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

18. 根据权利要求 16 所述的方法，其特征在于，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的位置信息；

所述根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号，包括：

根据所述目标虚拟扬声器的位置信息确定所述目标虚拟扬声器的 HOA 系数；

对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

19. 根据权利要求 15 至 18 中任一项所述的方法，其特征在于，所述虚拟扬声器信号是根据第一虚拟扬声器信号和第二虚拟扬声器信号下混获得的下混信号，所述方法还包括：

解码所述码流以获得边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；

根据所述边信息和所述下混信号获得所述第一虚拟扬声器信号和所述第二虚拟扬声器信号；

相应的，所述根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号，包括：

根据所述目标虚拟扬声器的属性信息、所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得所述重建的场景音频信号。

20. 一种音频编码装置，其特征在于，包括：

获取模块，用于根据当前场景音频信号从预设的虚拟扬声器集合中选择出第一目标虚拟扬声器；

信号生成模块，用于根据所述当前场景音频信号和所述第一目标虚拟扬声器的属性信息生成第一虚拟扬声器信号；

编码模块，用于对所述第一虚拟扬声器信号进行编码，以得到码流。

21. 根据权利要求 20 所述的装置，其特征在于，所述获取模块，用于根据所述虚拟扬声器集合从所述当前场景音频信号中获取主要声场成分；根据所述主要声场成分从所述虚拟扬声器集合中选择出所述第一目标虚拟扬声器。

22. 根据权利要求 21 所述的装置，其特征在于，所述获取模块，用于根据所述主要声场成分从高阶立体混响 HOA 系数集合中选择出与所述主要声场成分对应的 HOA 系数，所述 HOA 系数集合中的 HOA 系数与所述虚拟扬声器集合中的虚拟扬声器一一对应；确定所述虚拟扬声器集合中与所述主要声场成分对应的 HOA 系数对应的虚拟扬声器为所述第一目标虚拟扬声器。

23. 根据权利要求 21 所述的装置，其特征在于，所述获取模块，用于根据所述主要声场成分获取所述第一目标虚拟扬声器的配置参数；根据所述第一目标虚拟扬声器的配置参数生成所述第一目标虚拟扬声器对应的 HOA 系数；确定所述虚拟扬声器集合中所述第一目标虚拟扬声器对应的 HOA 系数对应的虚拟扬声器为所述目标虚拟扬声器。

24. 根据权利要求 23 所述的装置，其特征在于，所述获取模块，用于根据音频编码器的配置信息确定所述虚拟扬声器集合中的多个虚拟扬声器的配置参数；根据所述主要声场成分从所述多个虚拟扬声器的配置参数中选择出所述第一目标虚拟扬声器的配置参数。

25. 根据权利要求 23 或 24 所述的装置，其特征在于，所述第一目标虚拟扬声器的配置参数包括：所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息；

所述获取模块，用于根据所述第一目标虚拟扬声器的位置信息和 HOA 阶数信息确定所述第一目标虚拟扬声器对应的 HOA 系数。

26. 根据权利要求 20 至 25 中任一项所述的装置，其特征在于，所述编码模块，还用于对所述第一目标虚拟扬声器的属性信息进行编码，并写入所述码流。

27. 根据权利要求 20 至 26 中任一项所述的装置，其特征在于，所述当前场景音频信号，包括：待编码 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一目标虚拟扬声器的 HOA 系数；

所述信号生成模块，用于对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

28. 根据权利要求 20 至 26 中任一项所述的装置，其特征在于，所述当前场景音频信号包括：待编码高阶立体混响 HOA 信号；所述第一目标虚拟扬声器的属性信息包括所述第一

目标虚拟扬声器的位置信息；

所述信号生成模块，用于根据所述第一目标虚拟扬声器的位置信息获取所述第一目标虚拟扬声器对应的 HOA 系数；对所述待编码 HOA 信号和所述 HOA 系数进行线性组合，以得到所述第一虚拟扬声器信号。

29. 根据权利要求 20 至 28 中任一项所述的装置，其特征在于，

所述获取模块，用于根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

所述信号生成模块，用于根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

所述编码模块，用于对所述第二虚拟扬声器信号进行编码，并写入所述码流。

30. 根据权利要求 29 所述的装置，其特征在于，

所述信号生成模块，用于对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应地，所述编码模块，用于对所述对齐后的第二虚拟扬声器信号进行编码；

相应地，所述编码模块，用于对所述对齐后的第一虚拟扬声器信号进行编码。

31. 根据权利要求 20 至 28 中任一项所述的装置，其特征在于，

所述获取模块，用于根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器；

所述信号生成模块，用于根据所述当前场景音频信号和所述第二目标虚拟扬声器的属性信息生成第二虚拟扬声器信号；

相应地，所述编码模块，用于根据所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得下混信号和边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；对所述下混信号以及所述边信息进行编码。

32. 根据权利要求 31 所述的装置，其特征在于，

所述信号生成模块，用于对所述第一虚拟扬声器信号和所述第二虚拟扬声器信号进行对齐处理，以得到对齐后的第一虚拟扬声器信号和对齐后的第二虚拟扬声器信号；

相应的，所述编码模块，用于根据所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号获得所述下混信号和所述边信息；

相应的，所述边信息用于指示所述对齐后的第一虚拟扬声器信号和所述对齐后的第二虚拟扬声器信号之间的关系。

33. 根据权利要求 20 至 32 任一项所述的装置，其特征在于，所述获取模块，用于在根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器前，根据编码速率和/或所述当前场景音频信号的信号类型信息确定是否需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器；若需要获取除所述第一目标虚拟扬声器以外的目标虚拟扬声器，才根据所述当前场景音频信号从所述虚拟扬声器集合中选择出第二目标虚拟扬声器。

34. 一种音频解码装置，其特征在于，包括：

接收模块，用于接收码流；

解码模块，用于解码所述码流以获得虚拟扬声器信号；

重建模块，用于根据目标虚拟扬声器的属性信息以及所述虚拟扬声器信号获得重建的场景音频信号。

35. 根据权利要求 34 所述的装置，其特征在于，所述解码模块，还用于解码所述码流以获得所述目标虚拟扬声器的属性信息。

36. 根据权利要求 35 所述的装置，其特征在于，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的高阶立体混响 HOA 系数；

所述重建模块，用于对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

37. 根据权利要求 35 所述的装置，其特征在于，所述目标虚拟扬声器的属性信息包括所述目标虚拟扬声器的位置信息；

所述重建模块，用于根据所述目标虚拟扬声器的位置信息确定所述目标虚拟扬声器的 HOA 系数；对所述虚拟扬声器信号和所述目标虚拟扬声器的 HOA 系数进行合成处理，以获得所述重建的场景音频信号。

38. 根据权利要求 34 至 37 中任一项所述的装置，其特征在于，所述虚拟扬声器信号是根据第一虚拟扬声器信号和第二虚拟扬声器信号下混获得的下混信号，所述装置还包括：信号补偿模块，其中，

所述解码模块，用于解码所述码流以获得边信息，所述边信息用于指示所述第一虚拟扬声器信号和所述第二虚拟扬声器信号之间的关系；

所述信号补偿模块，用于根据所述边信息和所述下混信号获得所述第一虚拟扬声器信号和所述第二虚拟扬声器信号；

相应的，所述重建模块，用于根据所述目标虚拟扬声器的属性信息、所述第一虚拟扬声器信号和所述第二虚拟扬声器信号获得所述重建的场景音频信号。

39. 一种音频编码装置，其特征在于，所述音频编码装置包括至少一个处理器，所述至少一个处理器用于与存储器耦合，读取并执行所述存储器中的指令，以实现如权利要求 1 至 14 中任一项所述的方法。

40. 根据权利要求 39 所述的音频编码装置，其特征在于，所述音频编码装置还包括：所述存储器。

41. 一种音频解码装置，其特征在于，所述音频解码装置包括至少一个处理器，所述至少一个处理器用于与存储器耦合，读取并执行所述存储器中的指令，以实现如权利要求 15 至 19 中任一项所述的方法。

42. 根据权利要求 41 所述的音频解码装置，其特征在于，所述音频解码装置还包括：所述存储器。

43. 一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行如权利要求 1 至 14、或者 15 至 19 中任意一项所述的方法。

44. 一种计算机可读存储介质，包括如权利要求 1 至 14 任一项所述的方法所生成的码流。