

一种伴奏的生成方法、设备及存储介质

技术领域

本申请涉及计算机应用技术领域，尤其涉及一种伴奏的生成方法、设备及存储介质。

背景技术

随着虚拟现实（Virtual Reality）技术的发展，虚拟三维（Three-Dimensional，3D）音频技术也在逐渐优化。虚拟 3D 音频技术可以营造立体动态效果，将其应用到唱歌软件，能给用户带来身临其境的体验。目前，将虚拟 3D 音频技术应用在多人合唱的场景下时，现有的技术解决方案是将多路人声直接加权叠加，然而这种处理方法使得音效听感不够立体，导致用户体验不佳。

发明内容

本申请实施例提供了一种伴奏的生成方法、设备及存储介质，可全方位实现音频立体环绕的效果，提高用户体验。

一方面，本申请实施例提供了一种伴奏的生成方法，包括：

获取干声信号集合，干声信号集合中包括目标歌曲对应的 x 个干声信号， x 为大于 1 的整数；

基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号，其中， x 个干声信号对应 N 个虚拟三维空间声像位置， N 为大于 1 的整数， N 个虚拟三维空间声像位置不相同，且每一个虚拟三维空间声像位置被允许与 x 个干声信号中的一个或者多个干声信号对应；

对虚拟声音信号集合中的各个虚拟声音信号进行合并处理，得到合唱干声，该虚拟声音信号集合包括：在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号；

按照声效优化规则将合唱干声与目标歌曲的背景音乐进行音效合成处理，得到目标歌曲的伴奏。

一方面，本申请实施例提供了一种伴奏的播放处理方法，包括：

显示用户界面，用户界面用于接收对目标歌曲的选择指令；

若在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式，则获取目标歌曲对应的伴奏；

播放目标歌曲对应的伴奏；

伴奏是根据合唱干声和背景音乐生成的，合唱干声是根据干声信号集合中的多个干声信号生成的，干声信号集合中的多个干声信号对应多个不同的虚拟三维空间声像位置，干声信号集合是根据多个用户针对目标歌曲录入的干声信号得到的。

另一方面，本申请实施例提供了一种伴奏的生成装置，包括：

获取单元，用于获取干声信号集合，干声信号集合中包括目标歌曲对应的 x 个干声信号， x 为大于 1 的整数；基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号，其中， x 个干声信号对应 N 个虚拟三维空间声像位置， N 为大于 1 的整数， N 个虚拟三维空间声像位置不相同，且每一个虚拟三维空间声像位置被允许与 x 个干声信号中的一个或者多个干声信号对应。

处理单元，用于对虚拟声音信号集合中的各个虚拟声音信号进行合并处理，得到合唱干声，该虚拟声音信号集合包括：在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号；按照声效优化规则将合唱干声与目标歌曲的背景音乐进行音效合成处理，得到目标歌曲的伴奏。

另一方面，本申请实施例提供了一种伴奏的播放处理装置，包括：

获取单元，用于显示用户界面，用户界面用于接收对目标歌曲的选择指令；若在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式，则获取目标歌曲对应的伴奏。

处理单元，用于播放目标歌曲对应的伴奏；伴奏是根据合唱干声和背景音乐生成的，合唱干声是根

据干声信号集合中的多个干声信号生成的，干声信号集合中的多个干声信号对应多个不同的虚拟三维空间声像位置，干声信号集合是根据多个用户针对目标歌曲录入的干声信号得到的。

相应地，本申请实施例提供了一种电子设备，包括：存储器、处理器以及网络接口，上述处理器与上述存储器、上述网络接口相连，其中，上述网络接口用于提供网络通信功能，上述存储器用于存储程序代码，上述处理器用于调用程序代码，执行本申请实施例中的方法。

相应地，本申请实施例提供了一种计算机可读存储介质，包括：该计算机可读存储介质中存储有计算机程序，该计算机程序被处理器执行时，实现本申请实施例中的方法。

相应地，本申请实施例提供了一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中，计算机设备的处理器从计算机可读存储介质读取并执行该计算机指令，使得该计算机设备执行本申请实施例中的方法。

通过实施本申请实施例，一方面，可以获取干声信号集合中目标歌曲对应的各个干声信号在不同虚拟三维空间声像位置中的虚拟声音信号，然后将各个干声信号对应的虚拟声音信号进行合并处理后得到合唱干声，最终按照声效优化规则对合唱干声和目标歌曲的背景音乐进行音效合成处理，得到目标歌曲的伴奏；另一方面，可以接收用户对目标歌曲的选择指令，当接收到的选择指令指示目标歌曲的伴奏模式为合唱伴奏模式时，可以获取并播放该目标歌曲对应的伴奏。通过这种方式，可以全方位模拟干声信号在虚拟三维空间中的声像位置，实现音频立体环绕的效果，让用户在获取到相应的伴奏时，在听感上具有身临其境的感觉，获得沉浸式的体验。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图1是本申请实施例提供了一种伴奏的生成方法的应用场景示意图；

图2是本申请实施例提供了一种伴奏的生成方法的流程示意图；

图3a是本申请实施例提供了一种伴奏的生成方法中水平面、上方平面和下方平面的示意图；

图3b是本申请实施例提供了一种伴奏的生成方法中虚拟三维空间声像位置的示意图；

图3c是本申请实施例提供了一种伴奏的生成方法中以预设角度间隔对各个平面进行划分的示意图；

图4是本申请实施例提供的另一种伴奏的生成方法的流程示意图；

图5是本申请实施例提供了一种伴奏的生成方法中获取干声信号集中干声信号对应的双声道信号的流程示意图；

图6是本申请实施例提供了一种伴奏的播放处理方法的流程示意图；

图7a是本申请实施例提供了一种伴奏的播放处理方法中获取目标歌曲对应的伴奏的流程示意图；

图7b是本申请实施例提供了一种伴奏的播放处理方法中显示第一单句界面的示意图；

图7c是本申请实施例提供了一种伴奏的播放处理方法中显示第二单句界面的示意图；

图8a是本申请实施例提供了一种伴奏的生成装置的结构示意图；

图8b是本申请实施例提供了一种伴奏的播放处理装置的结构示意图；

图9是本申请实施例提供了一种电子设备的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

在对本申请实施例进行进一步详细说明之前，对本申请实施例中涉及的名词和术语进行说明，本申

请实施例所涉及的名词和术语适用于如下的解释。

1) 干声信号：本申请实施例中的干声信号是指没有伴奏音乐的纯人声信号，该干声信号为单声道声音信号，即不包括任何方向信息。

2) 双声道信号：双声道就是指有两个声音通道，其原理是人们听到声音时可以根据左耳和右耳对声音相位差来判断声源的具体位置。本申请实施例中的双声道信号即指左声道声音信号和右声道声音信号。

3) 头相关传递函数 (Head Related Transfer Functions, HRTF): HRTF 也可以称为双耳传输函数，描述了声波从声源到双耳的传输过程。HRTF 是一组滤波器，采用时域卷积等价于频域卷积的原理，可以根据声源位置信息对应的 HRTF 数据计算得到传输到双耳的虚拟声音信号。

本申请实施例提供了一种伴奏的生成方法、设备及存储介质，通过实施本申请实施例，可以获取由同一目标歌曲对应的多个干声信号构成的干声信号集合，并获取该干声信号集合中包括的各个干声信号在不同虚拟三维空间声像位置中所对应的虚拟声音信号，然后将各个干声信号对应的虚拟声音信号进行合并处理后得到合唱干声，最终按照声效优化规则对合唱干声和目标歌曲的背景音乐进行音效合成处理，得到目标歌曲的伴奏。通过这种方式，一方面，可以全方位模拟干声信号集合中包括的各个干声信号在虚拟三维空间中的声像位置，进而获得由各个干声信号在不同虚拟三维空间声像位置处对应的虚拟声音信号合并处理后的合唱干声，实现了音频立体环绕的效果，另一方面，可以按照声效优化规则将合唱干声和背景音乐进行音效合成处理后再得到伴奏，以此增强音频效果的沉浸感，总的来说，相较于直接对各个干声信号进行叠加的处理方式，本申请可以得到更加丰富的音频处理效果，提高了用户体验。

请参见图 1，图 1 是本申请实施例提供的一种伴奏的生成方法的应用场景示意图。如图 1 所示，该应用场景中可以包括智能设备 100，智能设备通过有线或者无线的方式与服务器 110 进行通信，服务器 110 连接数据库 120。

本申请实施例提供的伴奏的生成方法可以由诸如智能设备 100 等电子设备实现。例如，智能设备 100 在接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式时，可以获取目标歌曲对应的干声信号集合，并基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号，例如该虚拟声音信号可以为双声道信号，然后将各个干声信号对应的虚拟声音信号进行合并处理后得到合唱干声，再按照声效优化规则对合唱干声与目标歌曲的背景音乐进行音效合成处理，得到伴奏。作为示例，在图 1 的智能设备 100 中示出了“合唱伴奏”的选项，用户可以通过语音控制生成合唱伴奏模式的选择指令，也可以通过触发用户界面上显示的选择控件后生成合唱伴奏模式的选择指令。其中，干声信号集合可以是智能设备 100 预先存储在本地的，也可以是智能设备 100 从服务器 110 或数据库 120 中获取的。

本申请实施例提供的伴奏的生成方法也可以由诸如服务器 110 等电子设备实现。例如，服务器 110 在接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式时，可以获取目标歌曲对应的干声信号集合，并基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号，该虚拟声音信号例如可以是双声道信号，然后将各个干声信号对应的虚拟声音信号进行合并处理后得到合唱干声，再按照声效优化规则对合唱干声与目标歌曲的背景音乐进行音效合成处理，得到伴奏。其中，干声信号集合可以是服务器 110 预先存储在本地的，也可以是服务器 110 从数据库 120 中获取的，最终得到的伴奏可以存储在本地的，也可以存储至数据库 120，以在需要时调用。当然，服务器 110 也可以不用在接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式时时开始生成伴奏，服务器 110 可以在合适的时间，例如在服务器 110 负载较低时、或者服务器 110 接收到了目标歌曲的新的干声信号时、或者在接收到关于生成伴奏的管理操作时，开始执行本申请的伴奏的生成方法的相关步骤，生成伴奏。优选的，可以事先生成合唱版的伴奏，然后存储在服务器中，在生成大量歌曲的伴奏之后，用

户就可以通过智能设备 100，在用户界面上选择“合唱伴奏”等方式发出对目标歌曲的选择指令，这样一来服务器 110 就可以响应该选择指令，从生成的大量伴奏中找到目标歌曲的合唱伴奏，并将合唱伴奏下发给智能设备 100。

本申请实施例提供的伴奏的生成方法也可以由诸如智能设备 100 的电子设备和诸如服务器 110 的电子设备协同实现。例如，服务器 110 可以基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号，该虚拟声音信号例如可以是双声道信号，然后将各个干声信号对应的虚拟声音信号进行合并处理后得到合唱干声，再按照声效优化规则对合唱干声与目标歌曲的背景音乐进行音效合成处理，得到伴奏，并将得到伴奏下发至智能设备 100。

本申请实施例提供的伴奏的生成方法还可以由诸如智能设备 100 的电子设备和诸如服务器 110 的电子设备通过运行计算机程序来实现。例如，计算机程序可以是操作系统中的原生程序或软件模块，可以是本地应用程序（APP，Application），也可以是小程序，总之，该计算机程序可以是任意形式的应用程序、模块或插件，本申请实施例对此不做具体限定。

本申请实施例中所涉及的智能设备可以是个人计算机、笔记本电脑、智能手机、平板电脑、智能手表、智能语音交互设备、智能家电、车载终端和智能可穿戴设备等，但并不局限于此。服务器可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、内容分发网络（Content Delivery Network，CDN）以及大数据和人工智能平台等基础云计算服务的云服务器。智能设备以及服务器可以通过有线或无线通信方式进行直接或间接地连接，本申请实施例对此不做具体限定。

应当理解，图 1 中示出的干声信号和虚拟三维空间声像位置的数目仅仅是示意性的，根据实现需要，干声信号集合中可以具有任意数目的干声信号，虚拟三维空间中也可具有任意数目的虚拟三维声像位置。

进一步的，请参见图 2，图 2 是本申请实施例提供的一种伴奏的生成方法的流程示意图，本申请实施例的所述方法可以应用在电子设备上，该电子设备例如可以是智能手机、平板电脑、智能可穿戴设备、个人电脑等智能设备，也可以是服务器等等。该方法可以包括但不限于如下步骤：

S201：获取干声信号集合。

在本申请实施中，电子设备可以获取干声信号集合，该干声信号集合中包括目标歌曲对应的若干个干声信号。

在一个实施例中，干声信号集合可以从音频数据库中获取的，该音频数据库中包括多个用户演唱同一歌曲时录入的初始干声信号，需要说明的是，该音频数据库中的初始干声信号是经用户授权同意后录入的。电子设备可以根据初始干声信号的声音参数筛选出满足条件的干声信号，以构成干声信号集合。

在一个实施例中，电子设备可以根据音准特征参数和音质特征参数从初始干声信号集合中筛选满足条件的干声信号。其中，音准特征参数可以包括音调参数、节奏参数和韵律参数中的任意一种或多种，根据音准特征参数筛选出的满足条件的干声信号具有歌曲音调与节奏与伴奏旋律一致性高的特点；音质特征参数可以包括噪音参数、能量参数和速度参数中的任意一种或多种，根据音质特征参数筛选出的满足条件的干声信号具有音频清晰、音频能量适宜和音频速度均匀等特点。本申请实施例对满足条件的干声信号的筛选顺序不做限定，例如，电子设备可以先根据音准特征参数筛选满足条件的干声信号，再从满足预设音准特征参数条件的干声信号中筛选满足预设音质特征参数条件的干声信号，也可以先根据音质特征参数筛选满足条件的干声信号，再从满足预设音质特征参数条件的干声信号中筛选满足预设音频特征参数条件的干声信号。以这种方式从初始干声信号集合筛选得到的干声信号所构成的干声信号集合音准和音质俱佳。

S202 :基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号。

在本申请实施例中,电子设备可以模拟干声信号集合中各个干声信号在虚拟三维空间中不同的声像位置,再基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号,该虚拟声音信号例如可以是双声道信号。其中,N 个虚拟三维空间声像位置不相同,且每一个虚拟三维空间声像位置可以与一个或者多个干声信号对应。

在一个实施例中,N 个虚拟三维空间声像位置可以通过如下方式在虚拟三维空间中模拟得到:如图 3a 所示,虚拟三维空间中 x, y, z 轴的正方向分别对应人头的正前方,左侧方及头顶方向,将虚拟三维空间划分为三个平面:水平面 301、与水平面之间夹角为第一角度阈值的上方平面 302 和与水平面之间夹角为第二角度阈值的下方平面 303。如图 3b 所示,虚拟三维空间中的每个虚拟三维空间声像位置包括方位角和仰角,假设用 θ 表示虚拟三维空间声像位置的方位角,用 φ 表示虚拟三维空间声像位置的仰角,则每个虚拟三维空间声像位置可以用 (θ, φ) 表示。相应地,水平面 301 即为 0° 仰角对应的平面,上方平面即为第一角度阈值仰角对应的平面,该第一角度阈值可以为水平面以上的任意角度值,下方平面即为第二角度阈值仰角对应的平面,该第二角度阈值可以为水平面以下的任意角度值。例如,上方平面可以为 40° 仰角对应的平面,下方平面可以为 -40° 仰角对应的平面。其中,方位角 θ 可以用于描述虚拟三维空间声像位置在平面上依顺时针方向到目标方向线之间的夹角。进一步的,如图 3c 所示,在不同仰角对应的平面上以各自对应的预设角度为间隔进行划分后,可以得到多个虚拟三维空间声像位置。具体的,在水平面上以第一预设角度为间隔进行划分后,可以得到在水平面上的 n_1 个虚拟三维空间声像位置;以第二预设角度为间隔对上方平面进行划分后,可以得到在上方平面上的 n_2 个虚拟三维空间声像位置;以第三预设角度为间隔对下方平面进行划分后,可以得到下方平面上的 n_3 个虚拟三维空间声像位置。例如,假设第一预设角度为 10° ,第二预设角度和第三预设角度都为 15° ,那么以 10° 为间隔对水平面进行划分后将可以得到 36 个虚拟三维空间声像位置,以 15° 为间隔对上方平面进行划分后将可以得到 24 个虚拟三维空间声像位置,以 15° 为间隔对下方平面进行划分后也同样可以得到 24 个虚拟三维空间声像位置,如此合计将可以得到 84 个不同的虚拟三维空间声像位置。需要说明的是,本申请实施例中的第一预设角度、第二预设角度和第三预设角度可以是预设的任意角度值,上述三个预设角度的具体数值仅用于举例并不构成对本申请实施例的限定。通过这种方式,可以在虚拟三维空间中三个不同的平面内以不同间隔的方位角虚拟出多个虚拟三维空间声像位置,实现了声源的全方位沉浸感模拟。

在一个实施例中,每个虚拟三维空间声像位置处可以与一个干声信号对应,也可以与多个干声信号对应。电子设备可以获取每个虚拟三维空间声像位置处的一个或多个干声信号对应的虚拟声音信号,具体的,电子设备可以采用如下方案得到虚拟三维空间中各个干声信号对应的虚拟声音信号:获取干声信号对应的虚拟三维空间声像位置的方位角和仰角,根据该虚拟三维空间声像位置的方位角和仰角确定该虚拟三维空间声像位置对应的头部相关传递函数 HRTF,根据该虚拟三维空间声像位置的方位角和仰角与对应的 HRTF 数据可以计算得到虚拟空间声像位置处的干声信号对应的虚拟声音信号。例如,干声信号 X 对应的虚拟三维空间声像位置的方位角和仰角为 (θ, φ) ,该虚拟三维空间声像位置对应的 HRTF 数据表达式为:

$$\text{HRTF}(\theta, \varphi) \begin{cases} Y_L = X \cdot H_L \\ Y_R = X \cdot H_R \end{cases};$$

计算得到该虚拟三维空间声像位置处干声信号 X 对应的虚拟声音信号为双声道信号 Y_L 和 Y_R ,其中, Y_L 为左声道信号, Y_R 为右声道信号。

在一个实施例中,电子设备可以从干声信号集合中获取部分干声信号,例如可以随机获取或者按照新的筛选规则筛选出音准、音质更加出色的干声信号,对筛选的该部分干声信号分别进行延时处理操作,

得到该部分干声信号中各个干声信号对应的延时双声道信号。具体的, 在对一个干声信号进行延时处理操作时, 可以选取 8 对不同的时间参数, 需要说明的是, 8 对时间参数表示用于获取延时左声道信号的 8 个时间参数和用于获取延时右声道信号的 8 个时间参数, 总共将选取 16 个时间参数。例如, 可以从 21ms~79ms (根据一般房间脉冲响应取 80ms 作为混响时间选取) 范围内选取不等的 16 个参数作为时间参数, 也可以根据实际需要, 在其他合理范围内随机选取 16 (或者其他数值) 个不等的参数作为时间参数。通过这种方式, 可以模拟位于人头左耳处或者人头右耳处的干声信号, 让音频效果更加丰富。在一个实施例中, 关于选取的方式, 以及延时处理操作时的时间参数 (延时时长参数) 的设置, 可以通过一个界面进行调节, 以方便制作合唱音频的用户进行灵活配置。需要说明的是, 上述获取虚拟声音信号的步骤、获取延时左声道信号和延时右声道信号的步骤可以同时执行, 也可以先后执行, 本申请对此不做限定。

S203: 对虚拟声音信号集合中的各个虚拟声音信号进行合并处理, 得到合唱干声。

在本申请实施例中, 电子设备可以对虚拟声音信号集合中的各个虚拟声音信号进行合并处理, 以得到合唱干声。

在一个实施例中, 各个干声信号对应的虚拟声音信号进行合并处理可以通过归一化处理实现, 以达到将合并处理之后的虚拟声音信号的响度调整至 $[-1\text{dB}, 1\text{dB}]$ 的目的。其中, 进行合并处理时所针对的各个虚拟声音信号包括: 获取的 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的干声信号对应的虚拟声音信号, 和对干声信号集合中部分干声信号进行延时处理操作后得到的各个延时双声道信号。

S203: 按照声效优化规则对合唱干声与目标歌曲的背景音乐进行音效合成处理, 得到伴奏。

在本申请实施例中, 电子设备按照声效优化规则对合唱干声与目标歌曲的背景音乐进行音效合成处理后, 可以得到最终的伴奏。该声效优化规则例如可以是调整目标歌曲的背景音乐和上述获取到的多个干声信号对应的虚拟声音信号的声音参数, 该声音参数例如可以是响度、音色等等常见可调参数。

在一个实施例中, 电子设备在得到合唱干声后, 可以获取目标歌曲的背景音乐, 若得到的合唱干声与目标歌曲的背景音乐之间的能量关系不满足能量比条件, 则电子设备可以调整合唱干声与目标歌曲的背景音乐之间的能量关系。这里, 能量比条件可以设置为合唱干声的能量值与目标歌曲的背景音乐的能量值之间的比例的小于一个比例阈值, 也可以设置为合唱干声的响度比目标歌曲的背景音乐的响度低 3dB。通过这种方式, 可以避免合唱干声的能量大过目标歌曲的背景音乐, 使最终得到的伴奏更加和谐。

通过实施本申请实施例, 可以获得各个干声信号在虚拟三维空间中不同的虚拟三维空间声像位置处对应的虚拟声音信号, 然后将各个虚拟声音信号进行合并处理后得到合唱干声, 再按照声效优化规则对合唱干声与目标歌曲的背景音乐进行音效合成处理, 得到目标歌曲的伴奏, 以此实现了音频听感上立体环绕的效果, 并增强了音频效果的沉浸感, 使用户体验俱佳。

进一步的, 请参见图 4, 图 4 是本申请实施例提供的另一种伴奏的生成方法的流程示意图, 本申请实施例的所述方法可以应用在电子设备上, 该电子设备例如可以是智能手机、平板电脑、智能可穿戴设备、个人电脑、服务器等等。该方法可以包括但不限于如下步骤:

S401: 从音频数据库中获取初始干声信号集合。

在本申请实施例中, 电子设备可以从音频数据库中获取初始干声信号集合, 需要说明的是, 该音频数据库中的初始干声信号集合是经用户授权同意后录入的。

在一个实施例中, 音频数据库可以为独立设置的数据库, 也可以与电子设备集成在一起, 即音频数据库可以视为存储在电子设备内部。这里, 初始干声信号集合是指音频数据库中用户演唱同一首歌曲时授权同意后录入的原始干声信号所构成的集合。

S402: 根据各个初始干声信号的声音参数从初始干声信号集合中筛选出干声信号, 筛选出的干声信

号构成干声信号集合。

在本申请实施例中，电子设备可以根据各个初始干声信号的声音参数从初始干声信号集合中筛选出满足条件的干声信号，以缩小初始干声信号集合后构成干声信号集合。

在一个实施例中，初始干声信号的声音参数可以包括初始干声信号的音准特征参数和音质特征参数，音准特征参数可以包括音调参数、节奏参数和韵律参数中的任意一种或多种，音质特征参数可以包括噪音参数、能量参数和速度参数中的任意一种或多种。通过这种方式，可以将初始干声信号集合中听感嘈杂、出现跑调、音频时间过短、音频能量低和爆音等音频效果差的初始干声信号去除，以得到音准和音质俱佳的干声信号集合。

S403：获取 N 个虚拟三维空间声像位置每个虚拟三维空间声像位置对应的头部相关传递函数。

在本申请实施例中，电子设备可以获取虚拟三维空间中的 N 个虚拟三维空间声像位置，再根据 N 个虚拟三维空间声像位置获取每个虚拟三维空间声像位置对应的头部相关传递函数。

在一个实施例中，虚拟三维空间中各个虚拟三维空间声像位置对应的头部相关传递函数可以预先保存至头部相关传递函数数据库中，以供电子设备根据虚拟三维空间声像位置从头部相关传递函数数据库中调用对应的头部相关传递函数。

S404：通过目标虚拟三维空间声像位置对应的头部相关传递函数，对该目标虚拟三维空间声像位置对应的干声信号进行处理，得到目标虚拟三维空间声像位置处的虚拟声音信号。

在本申请实施例中，电子设备可以根据目标虚拟三维空间声像位置对应的头部相关传递函数对目标干声信号进行处理，以得到目标干声信号在目标虚拟三维空间声像位置处对应的虚拟声音信号，该目标虚拟三维空间声像位置可以是 N 个虚拟三维空间声像位置中的任意一个虚拟三维空间声像位置，该目标干声信号可以为干声信号集合中的任意一个干声信号。

在一个实施例中，目标虚拟三维空间声像位置对应的头部相关传递函数为该虚拟三维空间声像位置对应的 HRTF 数据。根据目标虚拟三维空间声像位置的方位角和仰角可以从已知的 HRTF 数据中确定与该目标虚拟三维空间声像位置对应的 HRTF 数据，然后电子设备可以将目标干声信号与目标虚拟三维空间位置对应的 HRTF 数据进行卷积，得到目标干声信号在目标虚拟三维空间声像位置处对应的虚拟声音信号。

S405：从干声信号集合包括的 x 个干声信号中获取 p 个干声信号。

在本申请实施例中，电子设备可以从干声信号集合包括的 x 个干声信号中随机获取 p 个干声信号。需要说明的是，S404 和 S405 可以同时执行，也可以先后执行，本申请并不限定。

S406：对 p 个干声信号中每个干声信号进行延时处理操作，得到 p 个干声信号中每个干声信号对应的延时左声道信号和延时右声道信号。

在本申请实施例中，电子设备可以将 p 个干声信号中的每一个干声信号进行 m_1 个时间参数的延时处理操作，得到 p 个干声信号中的每一个干声信号对应的 m_1 个延时干声信号，通过叠加每一个干声信号对应的 m_1 个延时干声信号得到 p 个干声信号中的每一个干声信号对应的延时左声道信号， m_1 为正整数；再将 p 个干声信号中的每一个干声信号进行 m_2 个时间参数的延时处理操作，得到 p 个干声信号中的每一个干声信号对应的 m_2 个延时干声信号，通过叠加每一个干声信号对应的 m_2 个延时干声信号得到 p 个干声信号中的每一个干声信号对应的延时右声道信号， m_2 为正整数。

在一个实施例中，电子设备可以将一个干声信号通过 16 个不同时间参数的延时器进行处理，以得到 16 个不同延时与衰减程度的干声信号，然后再将 16 个不同延时与衰减程度的干声信号平均分为两组，分别对每一组中不同延时与衰减程度的干声信号进行叠加，最终得到该干声信号对应的延时左声道信号和延时右声道信号。

在一个实施例中，得到该 p 个干声信号中每个干声信号对应的延时双声道信号之前，可以通过添加

低音增强和混响模拟模块的方式拉宽干声信号的声场，以降低通过延时处理得到的双声道信号中延时左声道信号和延时右声道信号之间的相关性。需要说明的是，S403 和 S404 获取虚拟声音信号的步骤、S405 和 S406 获取延时左声道信号和延时右声道信号的步骤可以同时执行，也可以先后执行，本申请对此不做限定，其中，S405 和 S406 为可选的两个步骤。

S407：对虚拟声音信号集合中的各个虚拟声音信号进行合并处理，得到合唱干声。

在本申请实施例中，电子设备可以对虚拟声音信号集合中的各个虚拟声音信号进行合并处理，以得到合唱干声。这里，虚拟声音信号集合包括电子设备通过模拟 N 个虚拟三维空间位置获取到的各个干声信号对应的虚拟声音信号和电子设备通过对干声信号集合中 p 个干声信号进行延时处理操作获取到的延时双声道信号。

在一个实施例中，虚拟声音信号集合中的各个虚拟声音信号为双声道信号，双声道信号包括左声道信号和右声道信号，对各个虚拟声音信号进行合并处理可以将左声道信号和右声道信号分开处理，左声道信号和右声道信号适用于同一个处理规则。这里，合并处理可以通过归一化处理实现，以使合并处理之后的双声道信号的响度为 $[-1\text{dB}, 1\text{dB}]$ 。例如，假设有 1000 个双声道信号，其中包括左声道信号 1000 个和右声道信号 1000 个，则将每个左声道信号分别进行归一化处理，再将 1000 个归一化处理之后的左声道信号的相加之和除以 1000，便可以得到合并处理后的左声道信号，同理，将每个右声道信号进行归一化处理之后，再将 1000 个归一化处理之后的右声道信号的相加之和除以 1000，便可以得到合并处理后的右声道信号。如此，便可以得到合唱干声。

在一个实施例中，得到的合唱干声与目标歌曲的背景音乐之间的能量关系可能满足能量比条件，也可能不满足能量比条件。若得到的合唱干声与背景音乐之间的能量关系满足能量比条件，则可以忽略步骤 S408；相应地，若得到的合唱干声与背景音乐之间的能量关系不满足能量比条件，则执行步骤 S408。

S408：获取目标歌曲的背景音乐，并调整合唱干声与背景音乐之间的能量关系。

在本申请实施例中，电子设备可以目标歌曲的背景音乐，并调整合唱干声与对应的背景音乐之间的能量关系，调整后的合唱干声与调整后的背景音乐之间的能量关系满足能量比条件。

在一个实施例中，合唱干声的能量可能过大，导致盖过背景音乐的音量。通过对合唱干声与背景音乐的调整，可以使调整后的合唱干声与调整后的背景音乐之间的能量关系满足能量比条件，以这种方式可以应对合唱干声的能量过大的情况。其中，能量比条件可以设置为合唱干声的能量值与背景音乐的音量值之间的比例的小于一个比例阈值，也可以设置为合唱干声的响度比背景音乐的响度低 3dB。

在一个实施例中，获取目标歌曲的背景音乐之后，可以按照上述 S202 基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号的详细描述，将背景音乐也做相同处理，以得到具有相似效果的合唱干声与背景音乐，从而达到更加和谐统一的听觉体验。

S409：对合唱干声在预设频段处进行频谱均衡处理。

在本申请实施例中，电子设备可以将合唱干声在预设频段进行频谱均衡处理。

在一个实施例中，电子设备可以通过在预设频段添加频谱陷波处理，以达到频谱均衡的目的，例如，电子设备可以在 4kHz 附近添加 6dB 左右的频谱陷波处理。通过这种方式，可以使合唱干声的听感更加自然，防止出现由于频谱不协调而产生的高频电流声。

S410：获取背景音乐的响度。

在本申请实施例中，电子设备可以获取背景音乐的响度。

S411：若响度小于响度阈值，则将调整后背景音乐的响度提升至响度阈值。

在本申请实施例中，背景音乐的响度小于响度阈值，则电子设备可以将背景音乐的响度提升至响度阈值。例如，该响度阈值可以设置为 -14dB ，若背景音乐的响度小于 -14dB ，则电子设备可以将其上调至 -14dB 。

S412：得到伴奏。

在本申请实施例中，电子设备将合唱干声和背景音乐进行叠加，便可以得到最终的伴奏。需要说明的是，该伴奏可以是根据 S408~S411 中任意一个步骤或多个步骤组合得到的，并且，在一个实施例中，S408~S411 可以根据实际需要选择性执行，例如，有可能存在合唱干声与背景音乐之间的能量关系不需要调整的情况，因此不执行 S408。同样，对合唱干声在预设频段处进行频谱均衡处理也为可选。又比如，S410 和 S411 两个步骤也可以不执行。在图 4 中仅仅指出了为了使得伴奏更加和谐自然、音质更佳所采用的技术方案，并且，在 S408 中进行能量关系调整、S409 频谱均衡调整、以及 S410 和 S411 体现的响度调整，这三方面的先后顺序本申请也不做限定。

在一个实施中，得到最终的伴奏之后可以将其存储至数据库中，以便于电子设备在收到相同歌曲的合唱请求时可以从数据库中直接获取对应的伴奏。

通过实施本申请实施例，可以获取通过模拟 N 个虚拟三维空间位置得到各个干声信号对应的虚拟声音信号，还可以获取通过对各个干声信号进行延时处理操作得到的各个干声信号对应的延时双声道信号，以此丰富了合唱干声。除此之外，通过调整合唱干声与背景音乐的能​​量关系，使最终得到的伴奏听感更加和谐自然，让用户可以很明显地感受到合唱时的空间感和沉浸感。

进一步的，请参见图 5，图 5 是本申请实施例提供的一种伴奏的生成方法中获取虚拟声音信号的流程图示意图，获取虚拟声音信号号包括：获取 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的干声信号对应的虚拟声音信号、和对 p 个干声信号中每个干声信号进行延时处理操作，得到 p 个干声信号中每个干声信号对应的延时双声道信号。

在本申请实施例中，获取到干声信号集合之后，可以获取 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的干声信号对应的虚拟声音信号，也可以对干声信号集合中的 p 个干声信号中每个干声信号进行延时处理操作，得到 p 个干声信号中每个干声信号对应的延时双声道信号。

作为举例，如图 5 所示，对干声信号集合中包括的干声信号 X 和干声信号 W 分别通过上述两种方式获取对应的虚拟声音信号，其中，干声信号 X 和干声信号 W 可以是干声信号集合中的任意干声信号。具体的，电子设备在获取到干声信号集合中的干声信号 X 之后，可以根据虚拟三维空间声像位置的方位角和仰角描述该虚拟三维空间声像位置的位置信息，即 (θ, φ) 。然后根据该虚拟三维空间声像位置的位置信息 (θ, φ) 可以找到该虚拟三维空间声像位置对应的头部相关传递函数 $HRTF(\theta, \varphi)$ ，将干声信号 X 与该虚拟三维空间声像位置对应的头部相关传递函数 $HRTF(\theta, \varphi)$ 进行卷积运算后，便可以得到该虚拟三维空间声像位置处的干声信号对应的虚拟声音信号，该虚拟声音信号为双声道信号，其中包括左声道信号 Y_L 和右声道信号 Y_R 。通过这种方式得到的干声信号对应的虚拟声音信号可以增强用户的立体沉浸感。

除此之外，电子设备在获取到干声信号集合中的干声信号 W 之后，可以对干声信号 W 进行延时处理操作。作为举例，电子设备可以将干声信号 W 通过 $d_L(1)$ 、 $d_L(2)$ ，...， $d_L(8)$ 和 $d_R(1)$ 、 $d_R(2)$ ，...， $d_R(8)$ 总共 16 个不同时间参数的延时器进行延时处理，然后再将通过 $d_L(1)$ 、 $d_L(2)$ ，...， $d_L(8)$ 这 8 个延时器进行延时处理得到的 8 个干声信号进行叠加，以得到干声信号 W 对应的延时左声道信号 W_L ，将通过 $d_R(1)$ 、 $d_R(2)$ ，...， $d_R(8)$ 这 8 个延时器进行延时处理得到的 8 个干声信号进行叠加，以得到干声信号 W 对应的延时右声道信号 W_R 。通过这种方式得到的干声信号对应的延时双声道信号可模拟人头左耳处或者右耳处的双声道信号，丰富了用户的听感效果。

在一个实施例中，最终得到的虚拟声音信号集合包括了上述两种情况，即最后的虚拟声音信号集合为： $Z=\{ZL,ZR\}$ ， $ZL=Y_L+W_L$ ； $ZR=Y_R+W_R$ 。需要说明的是，上述获取虚拟声音信号的步骤、获取延时左声道信号和延时右声道信号的步骤可以同时执行，也可以先后执行，本申请对此不做限定。将干声信号集合中的干声信号采用上述两种不同的方式获取对应的虚拟声音信号，可以全方位展示合唱时的场景体

验，让音频效果更加丰富。

进一步的，请参见图 6，图 6 是本申请实施例提供的一种伴奏的播放处理方法。本申请实施例的所述方法可以应用在电子设备上，该电子设备例如可以是智能手机、平板电脑、智能可穿戴设备、个人电脑等智能设备，也可以是服务器等等。该方法可以包括但不限于如下步骤：

S601：显示用户界面。

在本申请实施例中，电子设备可以显示用户界面，该用户界面用于接收用户对目标歌曲的选择指令。

在一个实施例中，该选择指令可以包括对目标歌曲的伴奏模式的选择指令，目标歌曲的伴奏模式可以是合唱伴奏模式、原声伴奏模式和人工智能（Artificial Intelligence，AI）伴奏模式，但不限于此。

在一个实施例中，该选择指令可以是用户通过触发用户界面上显示的选择控件后生成的指令，也可以是用户通过语音控制该电子设备生成的选择指令，例如，用户对电子设备的语音控制可以是“请使用合唱伴奏模式播放”，这样，电子设备便可以生成指示对目标歌曲的伴奏模式为合唱伴奏模式的选择指令。

S602：若在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式，则获取目标歌曲对应的伴奏。

在本申请实施例中，电子设备若在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式，则可以获取该目标歌曲合唱伴奏模式下对应的伴奏。

在一个实施例中，用户界面上可以显示有对目标歌曲的伴奏模式的选择控件，该选择控件可以包括：合唱伴奏模式选择控件、原声伴奏模式选择控件。在获取目标歌曲对应的伴奏之前，电子设备可以检测是否获取到针对合唱伴奏模式选择控件的选择操作，若是，则确认在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式。

在一个实施例中，该合唱伴奏模式下对应的伴奏是根据合唱干声和背景音乐生成的。其中，合唱干声可以是根据虚拟声音信号集合生成的，该虚拟声音信号集合包括：根据获取到的干声信号集合生成的在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号，该干声信号集合中的多个干声信号可以对应多个不同的虚拟三维空间声像位置，且每一个虚拟三维空间声像位置可以与一个或者多个干声信号对应，该干声信号集合是根据多个用户针对目标歌曲录入的干声信号得到的，需要说明的是，用户针对目标歌曲的干声信号是经用户授权同意后录入的。具体的，合唱伴奏模式下对应的伴奏的生成方法可见上述图 2-图 5 所示实施例，在这里不再赘述。

S603：播放目标歌曲对应的伴奏。

在本申请实施例中，电子设备获取到目标歌曲合唱伴奏模式下对应的伴奏之后，可以向用户播放该伴奏。

在一个实施中，目标歌曲对应的伴奏可以应用于 K 歌的场景，用户可以在播放该伴奏时进行演唱，在经过用户授权同意的情况下，电子设备可以将采集到用户歌声并与目标歌曲对应的伴奏融合后再进行播放，让用户有身处演唱会现场般的独特体验。

在一个实施例中，如图 7a 所示，电子设备获取目标歌曲对应的伴奏可以包括但不限于如下步骤：

S701：向服务器发送伴奏请求。

在本申请实施例中，电子设备可以向服务器发送伴奏请求，该伴奏请求中可以包括目标歌曲的标识信息。

在一个实施例中，目标歌曲的标识信息是用于标识该目标歌曲唯一的信息，例如，该标识信息可以是目标歌曲的歌曲名称。

S702：接收服务器响应伴奏请求返回的合唱干声和背景音乐。

在本申请实施例中，电子设备可以接收服务器响应针对目标歌曲的伴奏请求返回的合唱干声和背景

音乐。

在一个实施例中，服务器可以分别返回合唱干声和背景音乐，也可以将合唱干声和背景音乐合并之后再返回，具体返回方式可以根据用户的设定进行选择。

S703：从合唱干声中确定目标合唱干声片段。

在本申请实施例中，电子设备可以根据返回的合唱干声确定目标合唱干声片段。

在一个实施例中，电子设备可以显示第一单句界面，如图 7b 所示，该第一单句界面按照合唱干声的时间播放节点顺序显示了合唱干声对应的文本数据中的各个单句。用户可以根据第一单句界面显示的各个单句，选择目标合唱干声片段。

在一个实施例中，该目标合唱干声片段可以是合唱干声中的部分单句构成的，也可以是合唱干声中的全部单句构成的，具体可以由用户的选择操作决定。

S704：根据目标合唱干声片段对应的合唱干声和背景音乐，得到目标歌曲对应的伴奏。

在本申请实施例中，电子设备根据用户选择的目标合唱干声片段对应的合唱干声和背景音乐，可以得到目标歌曲对应的伴奏。

在一个实施例中，电子设备可以显示第二单句界面，如图 7c 所示，该第二单句界面可以在该目标歌曲对应的伴奏播放过程中显示，并按照该伴奏的时间播放节点顺序显示该伴奏对应的文本数据中的各个单句。

在一个实施例中，电子设备还可以检测播放过程中是否获取到针对伴奏中合唱干声的静音选择操作，若接收到用户针对伴奏中的合唱干声的静音选择操作，则可以在当前时间播放节点取消伴奏中合唱干声的播放，仅保留伴奏中背景音乐的播放。

通过实施本申请实施例，一方面，可以接收用户对目标歌曲的选择指令，当用户的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式时，可以获取该目标歌曲对应的伴奏，并播放该伴奏；另一方面，目标歌曲在合唱伴奏模式下的伴奏是根据合唱干声和背景音乐生成的，可以从合唱干声中确定目标合唱干声片段，并根据目标合唱干声片段对应的合唱干声和背景音乐生成目标歌曲对应的伴奏。通过这种方式，可以在播放目标歌曲对应的伴奏时，让用户有身处演唱会现场的体验，且在听感上具有身临其境的感觉，除此之外，用户还可以灵活选择伴奏中的合唱干声，增强了伴奏的趣味性，提高了用户体验。

进一步的，请参见图 8a，图 8a 是本申请实施例提供的一种伴奏的生成装置的结构示意图，本申请实施例的所述装置可以应用在电子设备上，该电子设备例如可以是智能手机、平板电脑、智能可穿戴设备、个人电脑、服务器等等，在一个实施例中，如图 8a 所示，该伴奏的生成装置 80，可以包括：

获取单元 801，用于获取干声信号集合，干声信号集合中包括目标歌曲对应的 x 个干声信号， x 为大于 1 的整数；基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号，其中， x 个干声信号对应 N 个虚拟三维空间声像位置， N 为大于 1 的整数， N 个虚拟三维空间声像位置不相同，且每一个虚拟三维空间声像位置被允许与 x 个干声信号中的一个或者多个干声信号对应。

处理单元 802，用于对虚拟声音信号集合中的各个虚拟声音信号进行合并处理，得到合唱干声，该虚拟声音信号集合包括：在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号；按照声效优化规则将合唱干声与目标歌曲的背景音乐进行音效合成处理，得到目标歌曲的伴奏。

在一个实施例中，获取单元 801 还可以用于从音频数据库中获取初始干声信号集合，音频数据库中包括多个用户演唱同一歌曲时录入的初始干声信号；处理单元 802 还可以用于根据各个初始干声信号的声音参数从初始干声信号集合中筛选出干声信号，筛选出的干声信号构成干声信号集合。

在一个实施例中，干声信号集合中包括：根据音准特征参数和音质特征参数从初始干声信号集合中筛选出的干声信号；音准特征参数包括音调参数、节奏参数和韵律参数中的任意一种或多种；音质特征

参数包括噪音参数、能量参数和速度参数中的任意一种或多种。

在一个实施例中， N 个虚拟三维空间声像位置包括；在水平面上以第一预设角度为间隔对水平面进行划分后，得到的在水平面上的 n_1 个虚拟三维空间声像位置；在上方平面上以第二预设角度为间隔对上方平面进行划分后，得到的在上方平面上的 n_2 个虚拟三维空间声像位置；上方平面与水平面之间的夹角为第一角度阈值；在下方平面上以第三预设角度为间隔对下方平面进行划分后，得到的在下方平面上的 n_3 个虚拟三维空间声像位置；下方平面与水平面之间的夹角为第二角度阈值；其中， n_1 、 n_2 和 n_3 为正整数且和值等于 N 。

在一个实施例中，获取单元 801 还可以用于获取 N 个虚拟三维空间声像位置中每个虚拟三维空间声像位置对应的头部相关传递函数；处理单元 802 还可用于通过目标虚拟三维空间声像位置对应的头部相关传递函数，对该目标虚拟三维空间声像位置对应的干声信号进行处理，得到目标虚拟三维空间声像位置处的虚拟声音信号；该目标虚拟三维空间声像位置处的虚拟声音信号为双声道信号；该目标虚拟三维空间声像位置为 N 个虚拟三维空间声像位置中的任一虚拟三维空间声像位置。

在一个实施例中，虚拟声音信号集合还包括： p 个干声信号中的每一个干声信号对应的延时左声道信号和延时右声道信号；获取单元 801 还可以用于从干声信号集合包括的 x 个干声信号中获取 p 个干声信号， p 为正整数且小于或者等于 x ；处理单元 802 还可以用于将 p 个干声信号中的每一个干声信号进行 m_1 个时间参数的延时处理操作，得到 p 个干声信号中的每一个干声信号对应的 m_1 个延时干声信号，通过叠加每一个干声信号对应的 m_1 个延时干声信号得到 p 个干声信号中的每一个干声信号对应的延时左声道信号， m_1 为正整数；将 p 个干声信号中的每一个干声信号进行 m_2 个时间参数的延时处理操作，得到 p 个干声信号中的每一个干声信号对应的 m_2 个延时干声信号，通过叠加每一个干声信号对应的 m_2 个延时干声信号得到 p 个干声信号中的每一个干声信号对应的延时右声道信号， m_2 为正整数。

在一个实施例中，获取单元 801 还可以用于获取目标歌曲的背景音乐，处理单元 802 还可以用于调整合唱干声与背景音乐之间的能量关系，调整后的合唱干声与调整后的背景音乐之间的能量关系满足能量比条件；伴奏是根据调整后的合唱干声与背景音乐得到伴奏得到的。

在一个实施例中，处理单元 802 还可以用于对合唱干声在预设频段处进行频谱均衡处理；获取单元 801 还可以用于获取背景音乐的响度；处理单元 802 还可以用于若背景音乐的响度小于响度阈值，则将背景音乐的响度提升至响度阈值；伴奏是根据经频谱均衡处理后的合唱干声和经响度处理后的背景音乐得到的。

需要说明的是，图 8a 对应的实施例中未提及的内容以及各个步骤的具体实现方式可参见图 2-图 5 所示实施例以及前述内容，这里不再赘述。

进一步的，请参见图 8b，图 8b 是本申请实施例提供的一种伴奏的播放处理装置的结构示意图，本申请实施例的所述装置可以应用在电子设备上，该电子设备例如可以是智能手机、平板电脑、智能可穿戴设备、个人电脑、服务器等等，在一个实施例中，如图 8b 所示，该伴奏的播放处理装置 81，可以包括：

获取单元 811，用于显示用户界面，用户界面用于接收对目标歌曲的选择指令；若在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式，则获取目标歌曲对应的伴奏。

处理单元 812，用于播放目标歌曲对应的伴奏；伴奏是根据合唱干声和背景音乐生成的，合唱干声是根据干声信号集合中的多个干声信号生成的，干声信号集合中的多个干声信号对应多个不同的虚拟三维空间声像位置，干声信号集合是根据多个用户针对目标歌曲录入的干声信号得到的。

在一个实施例中，合唱干声是根据虚拟声音信号集合生成的，虚拟声音信号集合包括：根据获取到的干声信号集合生成的在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号；其中，干声信号集合中的多个干声信号对应 N 个虚拟三维空间声像位置， N 为大于 1 的整数， N 个

虚拟三维空间声像位置不相同,且每一个虚拟三维空间声像位置被允许与一个或者多个干声信号对应。

在一个实施例中,用户界面上显示有对目标歌曲的伴奏模式的选择控件,伴奏模式的选择控件包括:合唱伴奏模式选择控件、原声伴奏模式选择控件;获取目标歌曲对应的伴奏之前,处理单元 812 还可以用于检测是否获取到针对合唱伴奏模式选择控件的选择操作;若是,则确认在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式。

在一个实施例中,处理单元 812 还可以用于向服务器发送伴奏请求,该伴奏请求中包括目标歌曲的标识信息;获取单元 811 还可以用于接收服务器响应伴奏请求返回的合唱干声和背景音乐;处理单元 812 还可以用于从合唱干声中确定目标合唱干声片段;根据目标合唱干声片段对应的合唱干声和背景音乐,得到目标歌曲对应的伴奏。

在一个实施例中,处理单元 812 还可以用于显示第一单句界面,按照合唱干声的时间播放节点顺序显示合唱干声对应的文本数据中的各个单句;目标合唱干声片段是基于在第一单句界面上的单句选择操作来确定的。

在一个实施例中,处理单元 812 还可以用于显示第二单句界面,按照伴奏的时间播放节点顺序显示伴奏对应的文本数据中的各个单句;检测播放过程中是否获取到针对伴奏中合唱干声的静音选择操作;若是,则在当前时间播放节点取消合唱干声的播放。

需要说明的是,图 8b 对应的实施例中未提及的内容以及各个步骤的具体实现方式可参见图 2-图 7c 所示实施例以及前述内容,这里不再赘述。

进一步的,请参见图 9,图 9 是本申请实施例提供的一种电子设备的结构示意图。该电子设备可以包括:网络接口 901、存储器 902 和处理器 903,网络接口 901、存储器 902 和处理器 903 通过一条或多条通信总线连接,通信总线用于实现这些组件之间的连接通信。网络接口 901 可以包括标准的有线接口、无线接口(如 WIFI 接口)。存储器 902 可以包括易失性存储器(volatile memory),例如随机存取存储器(random-access memory, RAM);存储器 902 也可以包括非易失性存储器(non-volatile memory),例如快闪存储器(flash memory),固态硬盘(solid-state drive, SSD)等;存储器 902 还可以包括上述种类的存储器的组合。处理器 903 可以是中央处理器(central processing unit, CPU)。处理器 903 还可以进一步包括硬件芯片。上述硬件芯片可以是专用集成电路(application-specific integrated circuit, ASIC),可编程逻辑器件(programmable logic device, PLD)等。上述 PLD 可以是现场可编程逻辑门阵列(field-programmable gate array, FPGA),通用阵列逻辑(generic array logic, GAL)等。

可选的,存储器 902 还用于存储程序指令,处理器 903 还可调用该程序指令,以实现:

获取干声信号集合,干声信号集合中包括目标歌曲对应的 x 个干声信号, x 为大于 1 的整数;

基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号,其中, x 个干声信号对应 N 个虚拟三维空间声像位置, N 为大于 1 的整数, N 个虚拟三维空间声像位置不相同,且每一个虚拟三维空间声像位置被允许与 x 个干声信号中的一个或者多个干声信号对应;

对虚拟声音信号集合中的各个虚拟声音信号进行合并处理,得到合唱干声,该虚拟声音信号集合包括:在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号;

按照声效优化规则将合唱干声与目标歌曲的背景音乐进行音效合成处理,得到目标歌曲的伴奏。

在一个实施例中,处理器 903 还可调用该程序指令,以实现:从音频数据库中获取初始干声信号集合,音频数据库中包括多个用户演唱同一歌曲时录入的初始干声信号;根据各个初始干声信号的声音参数从初始干声信号集合中筛选出干声信号,筛选出的干声信号构成干声信号集合。

在一个实施例中,干声信号集合中包括:根据音准特征参数和音质特征参数从初始干声信号集合中筛选出的干声信号;音准特征参数包括音调参数、节奏参数和韵律参数中的任意一种或多种;音质特征参数包括噪音参数、能量参数和速度参数中的任意一种或多种。

在一个实施例中, N 个虚拟三维空间声像位置包括; 在水平面上以第一预设角度为间隔对水平面进行划分后, 得到的在水平面上的 n_1 个虚拟三维空间声像位置; 在上方平面上以第二预设角度为间隔对上方平面进行划分后, 得到的在上方平面上的 n_2 个虚拟三维空间声像位置; 上方平面与水平面之间的夹角为第一角度阈值; 在下方平面上以第三预设角度为间隔对下方平面进行划分后, 得到的在下方平面上的 n_3 个虚拟三维空间声像位置; 下方平面与水平面之间的夹角为第二角度阈值; 其中, n_1 、 n_2 和 n_3 为正整数且和值等于 N 。

在一个实施例中, 处理器 903 还可调用该程序指令, 以实现: 获取 N 个虚拟三维空间声像位置中每个虚拟三维空间声像位置对应的头部相关传递函数; 通过目标虚拟三维空间声像位置对应的头部相关传递函数, 对该目标虚拟三维空间声像位置对应的干声信号进行处理, 得到目标虚拟三维空间声像位置处的虚拟声音信号; 该目标虚拟三维空间声像位置处的虚拟声音信号为双声道信号; 该目标虚拟三维空间声像位置为 N 个虚拟三维空间声像位置中的任一虚拟三维空间声像位置。

在一个实施例中, 虚拟声音信号集合还包括: p 个干声信号中的每一个干声信号对应的延时左声道信号和延时右声道信号; 处理器 903 还可调用该程序指令, 以实现: 从干声信号集合包括的 x 个干声信号中获取 p 个干声信号, p 为正整数且小于或者等于 x ; 将 p 个干声信号中的每一个干声信号进行 m_1 个时间参数的延时处理操作, 得到 p 个干声信号中的每一个干声信号对应的 m_1 个延时干声信号, 通过叠加每一个干声信号对应的 m_1 个延时干声信号得到 p 个干声信号中的每一个干声信号对应的延时左声道信号, m_1 为正整数; 将 p 个干声信号中的每一个干声信号进行 m_2 个时间参数的延时处理操作, 得到 p 个干声信号中的每一个干声信号对应的 m_2 个延时干声信号, 通过叠加每一个干声信号对应的 m_2 个延时干声信号得到 p 个干声信号中的每一个干声信号对应的延时右声道信号, m_2 为正整数。

在一个实施例中, 处理器 903 还可调用该程序指令, 以实现: 获取目标歌曲的背景音乐, 并调整合唱干声与背景音乐之间的能量关系, 调整后的合唱干声与调整后的背景音乐之间的能量关系满足能量比条件; 伴奏是根据调整后的合唱干声与背景音乐得到伴奏得到的。

在一个实施例中, 处理器 903 还可调用该程序指令, 以实现: 对合唱干声在预设频段处进行频谱均衡处理; 获取背景音乐的响度; 若背景音乐的响度小于响度阈值, 则将背景音乐的响度提升至响度阈值; 伴奏是根据经频谱均衡处理后的合唱干声和经响度处理后的背景音乐得到的。

可选的, 存储器 902 还用于存储程序指令, 处理器 903 还可调用该程序指令, 以实现:

显示用户界面, 用户界面用于接收对目标歌曲的选择指令;

若在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式, 则获取目标歌曲对应的伴奏;

播放目标歌曲对应的伴奏; 伴奏是根据合唱干声和背景音乐生成的, 合唱干声是根据干声信号集合中的多个干声信号生成的, 干声信号集合中的多个干声信号对应多个不同的虚拟三维空间声像位置, 干声信号集合是根据多个用户针对目标歌曲录入的干声信号得到的。

在一个实施例中, 合唱干声是根据虚拟声音信号集合生成的, 虚拟声音信号集合包括: 根据获取到的干声信号集合生成的在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号; 其中, 干声信号集合中的多个干声信号对应 N 个虚拟三维空间声像位置, N 为大于 1 的整数, N 个虚拟三维空间声像位置不相同, 且每一个虚拟三维空间声像位置被允许与一个或者多个干声信号对应。

在一个实施例中, 用户界面上显示有对目标歌曲的伴奏模式的选择控件, 伴奏模式的选择控件包括: 合唱伴奏模式选择控件、原声伴奏模式选择控件; 获取目标歌曲对应的伴奏之前, 处理器 903 还可调用该程序指令, 以实现: 检测是否获取到针对合唱伴奏模式选择控件的选择操作; 若是, 则确认在用户界面接收到的选择指令指示对目标歌曲的伴奏模式为合唱伴奏模式。

在一个实施例中, 处理器 903 还可调用该程序指令, 以实现: 向服务器发送伴奏请求, 该伴奏请求

中包括目标歌曲的标识信息；接收服务器响应伴奏请求返回的合唱干声和背景音乐；从合唱干声中确定目标合唱干声片段；根据目标合唱干声片段对应的合唱干声和背景音乐，得到目标歌曲对应的伴奏。

在一个实施例中，处理器 903 还可调用该程序指令，以实现：显示第一单句界面，按照合唱干声的时间播放节点顺序显示合唱干声对应的文本数据中的各个单句；目标合唱干声片段是基于在第一单句界面上的单句选择操作来确定的。

在一个实施例中，处理器 903 还可调用该程序指令，以实现：显示第二单句界面，按照伴奏的时间播放节点顺序显示伴奏对应的文本数据中的各个单句；检测播放过程中是否获取到针对伴奏中合唱干声的静音选择操作；若是，则在当前时间播放节点取消合唱干声的播放。

应当理解，本申请实施例中所描述的电子设备 90 解决问题的原理与有益效果与本申请图 2-图 7c 所示实施例以及前述内容解决问题的原理和有益效果相似，为简洁描述，在这里不再赘述。

此外，本申请还提供了一种计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时，实现前述实施例提供的方法。

本申请实施例还提供了一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器执行该计算机指令，使得该计算机设备执行前述实施例提供的方法。

本申请实施例方法中的步骤可以根据实际需要进行顺序调整、合并和删减。

本申请实施例装置中的单元可以根据实际需要进行合并、划分和删减。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，该程序可存储于计算机可读存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，上述存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory, ROM)或随机存储记忆体(Random Access Memory, RAM)等。

以上所揭露的仅为本申请的部分实施例而已，当然不能以此来限定本申请之权利范围，本领域普通技术人员可以理解实现上述实施例的全部或部分流程，并依本申请权利要求所作的等同变化，仍属于本申请所涵盖的范围。

权利要求书

1、一种伴奏的生成方法，其特征在于，所述方法包括：

获取干声信号集合，所述干声信号集合中包括目标歌曲对应的 x 个干声信号，所述 x 为大于 1 的整数；

基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号，其中，所述 x 个干声信号对应 N 个虚拟三维空间声像位置， N 为大于 1 的整数， N 个虚拟三维空间声像位置不相同，且每一个虚拟三维空间声像位置被允许与所述 x 个干声信号中的一个或者多个干声信号对应；

对虚拟声音信号集合中的各个虚拟声音信号进行合并处理，得到合唱干声，所述虚拟声音信号集合包括：在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号；

按照声效优化规则将所述合唱干声与所述目标歌曲的背景音乐进行音效合成处理，得到所述目标歌曲的伴奏。

2、如权利要求 1 所述的方法，其特征在于，所述获取干声信号集合，包括：

从音频数据库中获取初始干声信号集合，所述音频数据库中包括多个用户演唱目标歌曲时录入的初始干声信号；

根据各个初始干声信号的声音参数从所述初始干声信号集合中筛选出 x 个干声信号构成所述干声信号集合。

3、如权利要求 2 所述的方法，其特征在于，所述声音参数包括：音准特征参数和音质特征参数；

其中，所述音准特征参数包括音调参数、节奏参数和韵律参数中的任意一种或多种；所述音质特征参数包括噪音参数、能量参数和速度参数中的任意一种或多种。

4、如权利要求 1 所述方法，其特征在于， N 个虚拟三维空间声像位置包括：

在水平面上以第一预设角度为间隔对所述水平面进行划分后，得到的在所述水平面上的 n_1 个虚拟三维空间声像位置；

在上方平面上以第二预设角度为间隔对所述上方平面进行划分后，得到的在所述上方平面上的 n_2 个虚拟三维空间声像位置；所述上方平面与所述水平面之间的夹角为第一角度阈值；

在下方平面上以第三预设角度为间隔对所述下方平面进行划分后，得到的在所述下方平面上的 n_3 个虚拟三维空间声像位置；所述下方平面与所述水平面之间的夹角为第二角度阈值；

其中，所述 n_1 、所述 n_2 和所述 n_3 为正整数且和值等于所述 N 。

5、如权利要求 1-4 任一项所述方法，其特征在于，所述基于 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处对应的干声信号生成虚拟声音信号，包括：

获取 N 个虚拟三维空间声像位置中每个虚拟三维空间声像位置对应的头部相关传递函数；

通过目标虚拟三维空间声像位置对应的头部相关传递函数，对该目标虚拟三维空间声像位置对应的干声信号进行处理，得到所述目标虚拟三维空间声像位置处的虚拟声音信号；

该目标虚拟三维空间声像位置处的虚拟声音信号为双声道信号；

所述目标虚拟三维空间声像位置为所述 N 个虚拟三维空间声像位置中的任一虚拟三维空间声像位置。

6、如权利要求 1-4 任一项所述的方法，其特征在于，所述虚拟声音信号集合还包括： p 个干声信号中的每一个干声信号对应的延时左声道信号和延时右声道信号；

所述对虚拟声音信号集合中的各个虚拟声音信号进行合并处理，得到合唱干声之前，所述方法还包括：

从所述干声信号集合包括的 x 个干声信号中获取 p 个干声信号, 所述 p 为正整数且小于或者等于所述 x ;

将所述 p 个干声信号中的每一个干声信号进行 $m1$ 个时间参数的延时处理操作, 得到 p 个干声信号中的每一个干声信号对应的 $m1$ 个延时干声信号, 通过叠加所述每一个干声信号对应的 $m1$ 个延时干声信号得到 p 个干声信号中的每一个干声信号对应的延时左声道信号, 所述 $m1$ 为正整数;

将所述 p 个干声信号中的每一个干声信号进行 $m2$ 个时间参数的延时处理操作, 得到 p 个干声信号中的每一个干声信号对应的 $m2$ 个延时干声信号, 通过叠加所述每一个干声信号对应的 $m2$ 个延时干声信号得到 p 个干声信号中的每一个干声信号对应的延时右声道信号, 所述 $m2$ 为正整数。

7、如权利要求 1-4 任一项所述的方法, 其特征在于, 所述按照声效优化规则将所述合唱干声与所述目标歌曲的背景音乐进行音效合成处理, 得到所述目标歌曲的伴奏, 包括:

获取所述目标歌曲的背景音乐, 并调整所述合唱干声与所述背景音乐之间的能量关系, 调整后的合唱干声与调整后的背景音乐之间的能量关系满足能量比条件;

所述伴奏是根据调整后的合唱干声与背景音乐得到的。

8、如权利要求 1-4 任一项所述的方法, 其特征在于, 所述按照声效优化规则将所述合唱干声与所述目标歌曲的背景音乐进行音效合成处理, 得到所述目标歌曲的伴奏, 包括:

对所述合唱干声在预设频段处进行频谱均衡处理;

获取所述背景音乐的响度;

若所述背景音乐的响度小于响度阈值, 则将所述背景音乐的响度提升至响度阈值;

所述伴奏是根据经频谱均衡处理后的合唱干声和经响度处理后的背景音乐得到的。

9、一种伴奏的播放处理方法, 其特征在于, 包括:

显示用户界面, 所述用户界面用于接收对目标歌曲的选择指令;

若在所述用户界面接收到的选择指令指示对所述目标歌曲的伴奏模式为合唱伴奏模式, 则获取所述目标歌曲对应的伴奏;

播放所述目标歌曲对应的伴奏;

所述伴奏是根据合唱干声和背景音乐生成的, 所述合唱干声是根据干声信号集合中的多个干声信号生成的, 所述干声信号集合中的多个干声信号对应多个不同的虚拟三维空间声像位置, 所述干声信号集合是根据多个用户针对所述目标歌曲录入的干声信号得到的。

10、如权利要求 9 所述的方法, 其特征在于, 所述合唱干声是根据虚拟声音信号集合生成的, 所述虚拟声音信号集合包括: 根据获取到的干声信号集合生成的在 N 个虚拟三维空间声像位置中的每个虚拟三维空间声像位置处的虚拟声音信号;

其中, 所述干声信号集合中的多个干声信号对应 N 个虚拟三维空间声像位置, N 为大于 1 的整数, N 个虚拟三维空间声像位置不相同, 且每一个虚拟三维空间声像位置被允许与一个或者多个干声信号对应。

11、如权利要求 9 或 10 所述的方法, 其特征在于, 所述用户界面上显示有对目标歌曲的伴奏模式的选择控件, 所述伴奏模式的选择控件包括: 合唱伴奏模式选择控件、原声伴奏模式选择控件; 所述获取所述目标歌曲对应的伴奏之前, 还包括:

检测是否获取到针对所述合唱伴奏模式选择控件的选择操作;

若是, 则确认在所述用户界面接收到的选择指令指示对所述目标歌曲的伴奏模式为合唱伴奏模式。

12、如权利要求 9 所述的方法, 其特征在于, 所述获取所述目标歌曲对应的伴奏, 包括:

向服务器发送伴奏请求, 所述伴奏请求中包括所述目标歌曲的标识信息;

接收所述服务器响应所述伴奏请求返回的所述合唱干声和所述背景音乐;

从所述合唱干声中确定目标合唱干声片段；

根据所述目标合唱干声片段对应的合唱干声和所述背景音乐，得到所述目标歌曲对应的伴奏。

13、如权利要求 12 所述的方法，其特征在于，所述从合唱干声中确定目标合唱干声片段之前，所述方法还包括：

显示第一单句界面，按照所述合唱干声的时间播放节点顺序显示所述合唱干声对应的文本数据中的各个单句；

所述目标合唱干声片段是基于在所述第一单句界面上的单句选择操作来确定的。

14、如权利要求 9 或 12 所述的方法，其特征在于，所述播放所述目标歌曲对应的伴奏之后，所述方法包括：

显示第二单句界面，按照所述伴奏的时间播放节点顺序显示所述伴奏对应的文本数据中的各个单句；

检测播放过程中是否获取到针对所述伴奏中所述合唱干声的静音选择操作；

若是，则在当前时间播放节点取消所述合唱干声的播放。

15、一种电子设备，其特征在于，包括：存储器、处理器以及网络接口，所述处理器与所述存储器、所述网络接口相连，其中，所述网络接口用于提供网络通信功能，所述存储器用于存储程序代码，所述处理器用于调用所述程序代码，执行权利要求 1~14 任一项所述的方法。

16、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中存储有计算机程序，该计算机程序被处理器执行时，实现权利要求 1~14 任一项所述的方法。