

音视频处理方法、装置、终端及存储介质

本申请要求于 2018 年 12 月 21 日提交的申请号为 201811575178.3、发明名称为“音视频处理方法、装置、终端及存储介质”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

本申请实施例涉及多媒体技术领域，特别涉及一种音视频处理方法、装置、终端及存储介质。

背景技术

音视频处理方法包括对音频和/或视频进行处理的方法。

相关技术中，终端播放目标音频时显示该目标音频的音频播放页面，该音频播放页面包括音乐短片（Music Video，MV）入口，当终端检测到作用于 MV 入口的点击信号时，将当前的音频播放页面跳转显示为 MV 播放页面，在 MV 播放页面上从头开始播放该音频对应的 MV。

发明内容

本申请实施例提供了一种音视频处理方法、装置、终端及存储介质，可以解决相关技术中音频对应的 MV 是固定不可变的导致音视频播放效果差的问题。技术方案如下：

一方面，提供了一种音视频处理方法，用于终端中，所述方法包括：

显示目标音频的视频创建界面，所述视频创建界面中包括所述目标音频的 n 个音频片段和所述 n 个音频片段各自对应的视频录制入口，其中， $n \geq 2$ ；

在所述视频创建界面中接收作用于目标视频录制入口的触发信号，所述目标视频录制入口为与目标音频片段对应的视频录制入口；

根据所述触发信号获取所述目标音频片段对应的目标视频，所述目标视频为视频时长小于时长阈值的视频片段；

向所述服务器发送携带有所述目标视频的视频创建请求，所述视频创建请求用于指示在播放所述目标音频时播放所述目标视频的画面信息。

可选的，所述根据所述触发信号获取所述目标音频片段对应的目标视频，包括：

根据所述触发信号开启视频录制功能，进行视频录制得到所述目标视频；或者，

根据所述触发信号获取所述终端中存储的所述目标视频；或者，

根据所述触发信号向所述服务器发送视频获取请求，所述视频获取请求中携带有所述目标音频片段的片段标识，所述视频获取请求用于指示所述服务器反馈所述片段标识对应的至少一个视频；从接收到的所述服务器反馈的所述至少一个视频中确定所述目标视频。

可选的，所述显示目标音频的视频创建界面，包括：

显示所述目标音频的音频关联界面，所述音频关联界面中包括视频创建入口；

当接收到作用于视频创建入口的触发信号时，显示所述目标音频的视频创建界面。

可选的，所述目标音频片段包括至少两个所述目标音频片段，所述至少两个目标音频片段各自对应有所述目标视频；

所述向所述服务器发送携带有所述目标视频的视频创建请求之前，还包括：

将至少两个所述目标视频各自对应的画面信息进行拼接得到第一视频拼接画面；

所述向所述服务器发送携带有所述目标视频的视频创建请求，包括：

向所述服务器发送携带有所述第一视频拼接画面的第一视频创建请求，所述第一视频创建请求用于指示在播放所述目标音频时播放所述第一视频拼接画面。

可选的，所述方法还包括：

向所述服务器发送所述目标音频的音频播放请求，所述音频播放请求用于指示所述服务器获取所述 n 个音频片段各自对应的第一候选视频并反馈画面播放指令，所述第一候选视频包括所述目标视频；

根据接收到的所述画面播放指令，在播放所述目标音频的音频片段时播放所述 n 个音频片段各自对应的所述第一候选视频的画面信息。

可选的，所述方法还包括：

当接收到作用于第一音频片段的画面切换操作信号时，获取所述目标音频

当前的第一播放时间点；

向所述服务器发送画面切换请求，所述画面切换请求携带有所述目标音频的音频标识，所述画面切换请求用于指示所述服务器获取与所述音频标识对应的第二视频拼接画面；向所述终端发送第一画面切换指令；

根据接收到的所述第一画面切换指令，在所述目标音频的第一播放时间点上开始切换播放所述第二视频拼接画面。

可选的，所述画面切换请求中还携带有所述第一播放时间点，所述画面切换请求还用于指示所述服务器当不存在与所述音频标识对应的所述第二视频拼接画面时，获取与所述第一播放时间点所在的所述第一音频片段对应的第二候选视频；向所述终端发送第二画面切换指令；

所述方法还包括：

根据接收到的所述第二画面切换指令，在播放所述第一音频片段的过程中在所述目标音频的所述第一播放时间点上开始切换播放所述第二候选视频的画面信息。

可选的，所述方法还包括：

在播放所述目标音频的指定时间内容时，将第一候选视频的画面信息切换显示为所述第一候选视频的视频制作信息，所述第一候选视频为所述指定时间内容所在的音频片段对应的候选视频；

所述指定时间内容包括前奏内容、间奏内容和尾奏内容中的至少一种，所述视频制作信息包括所述第一候选视频对应的录制者信息、上传者信息和拼接者信息中的至少一种。

根据本申请实施例的另一方面，提供了一种音视频处理装置，用于终端中，所述装置包括：

显示模块，用于显示目标音频的视频创建界面，所述视频创建界面中包括所述目标音频的 n 个音频片段和所述 n 个音频片段各自对应的视频录制入口，其中， $n \geq 2$ ；

接收模块，用于在所述视频创建界面中接收作用于目标视频录制入口的触发信号，所述目标视频录制入口为与目标音频片段对应的视频录制入口；

获取模块，用于根据所述触发信号获取所述目标音频片段对应的目标视频，所述目标视频为视频时长小于时长阈值的视频片段；

发送模块，用于向所述服务器发送携带有所述目标视频的视频创建请求，所述视频创建请求用于指示在播放所述目标音频时播放所述目标视频的画面信息。

可选的，所述获取模块，还用于根据所述触发信号开启视频录制功能，进行视频录制得到所述目标视频；

或，

根据所述触发信号获取所述终端中存储的所述目标视频；

或，

根据所述触发信号向所述服务器发送视频获取请求，所述视频获取请求中携带有所述目标音频片段的片段标识，所述视频获取请求用于指示所述服务器反馈所述片段标识对应的至少一个视频；从接收到的所述服务器反馈的所述至少一个视频中确定所述目标视频。

可选的，所述显示模块，还用于显示所述目标音频的音频关联界面，所述音频关联界面中包括视频创建入口；

当接收到作用于视频创建入口的触发信号时，显示所述目标音频的视频创建界面。

可选的，所述目标音频片段包括至少两个所述目标音频片段，所述至少两个目标音频片段各自对应有所述目标视频；

所述装置还包括：

拼接模块，用于将至少两个所述目标视频各自对应的画面信息进行拼接得到第一视频拼接画面；

发送模块，还用于向所述服务器发送携带有所述第一视频拼接画面的第一视频创建请求，所述第一视频创建请求用于指示在播放所述目标音频时播放所述第一视频拼接画面。

可选的，所述装置还包括：

播放模块，用于向所述服务器发送所述目标音频的音频播放请求，所述音频播放请求用于指示所述服务器获取所述 n 个音频片段各自对应的第一候选视频并反馈画面播放指令，所述第一候选视频包括所述目标视频；根据接收到的所述画面播放指令，在播放所述目标音频的音频片段时播放所述 n 个音频片段各自对应的所述第一候选视频的画面信息。

可选的，所述装置还包括：

第一画面切换模块，用于当接收到作用于第一音频片段的画面切换操作信号时，获取所述目标音频当前的第一播放时间点；

向所述服务器发送画面切换请求，所述画面切换请求携带有所述目标音频的音频标识，所述画面切换请求用于指示所述服务器获取与所述音频标识对应的第二视频拼接画面；向所述终端发送第一画面切换指令；

根据接收到的所述第一画面切换指令，在所述目标音频的第一播放时间点上开始切换播放所述第二视频拼接画面。

可选的，所述画面切换请求中还携带有所述第一播放时间点，所述画面切换请求还用于指示所述服务器当不存在与所述音频标识对应的所述第二视频拼接画面时，获取与所述第一播放时间点所在的所述第一音频片段对应的第二候选视频；向所述终端发送第二画面切换指令；

所述装置还包括：

根据接收到的所述第二画面切换指令，在播放所述第一音频片段的过程中在所述目标音频的所述第一播放时间点上开始切换播放所述第二候选视频的画面信息。

可选的，该装置还包括：切换显示模块，用于在播放所述目标音频的指定时间内容时，将第一候选视频的画面信息切换显示为所述第一候选视频的视频制作信息，所述第一候选视频为所述指定时间内容所在的音频片段对应的候选视频；所述指定时间内容包括前奏内容、间奏内容和尾奏内容中的至少一种，所述视频制作信息包括所述第一候选视频对应的录制者信息、上传者信息和拼接者信息中的至少一种。

根据本申请实施例的另一方面，提供了一种终端，所述终端包括处理器和存储器；所述存储器存储有至少一条指令，所述至少一条指令用于被所述处理器执行以实现如上述本申请实施例任一所述的音视频处理方法。

根据本申请实施例的另一方面，提供了一种计算机可读存储介质，所述存储介质存储有至少一条指令，所述至少一条指令用于被处理器执行以实现如上述本申请实施例任一所述的音视频处理方法。

本实施例提供的音视频处理方法，通过显示目标音频的视频创建界面，在视频创建界面中接收作用于目标音频片段对应的视频录制入口的触发信号，根据触发信号获取目标音频片段对应的目标视频，向服务器发送携带有目标视频

的视频创建请求，使得终端可以针对一个音频中的一个或者多个音频片段上传相应的视频，进而使得后续在播放该音频的过程中能够同步播放音频片段对应的视频的画面信息，避免了相关技术中音频对应的 MV 是固定不可变的导致音视频播放效果差的问题，提高了音视频的播放效果。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例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 是本申请一个示例性实施例所提供的终端的结构示意图；

图 2 是本申请一个示例性实施例提供的音视频处理方法的流程图；

图 3 是本申请另一个示例性实施例提供的音视频处理方法的流程图；

图 4 是本申请另一个示例性实施例提供的音视频处理方法涉及的界面示意图；

图 5 是本申请另一个示例性实施例提供的音视频处理方法的流程图；

图 6 是本申请一个示例性实施例提供的音视频处理装置的结构示意图；

图 7 是本申请另一个示例性实施例提供的音视频处理装置的结构示意图；

图 8 是本申请一个示例性实施例提供的终端的结构框图；

图 9 是本申请一个示例性实施例提供的服务器的结构示意图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请实施方式作进一步地详细描述。

在本文中提及的“多个”是指两个或两个以上。“和/或”，描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。字符“/”一般表示前后关联对象是一种“或”的关系。

在对本申请实施例进行解释说明之前，先对本申请实施例的应用场景进行

说明。该音视频处理系统包括终端 110 和服务器 120。

终端 110 是具有视频上传功能的电子设备。可选的，终端 110 用于将目标视频上传至服务器 120 中。比如，终端 110 可以是手机、平板电脑、膝上型便携计算机和台式计算机等等。

可选的，终端 110 是具有视频录制功能的电子设备。终端 110 用于通过图像采集组件（如：摄像头）和音频采集组件（如：麦克风）录制目标视频，并将录制好的目标视频上传至服务器 120 中。

可选的，终端 110 中运行有音频应用，终端 110 用于显示音频应用中的目标音频的视频创建界面，该视频创建界面中包括目标音频的多个音频片段和多个音频片段各自对应的视频录制入口；在视频创建界面中接收作用于目标音频片段对应的目标视频录制入口的触发信号，根据触发信号获取目标音频片段对应的目标视频，向服务器 120 发送携带有该目标视频的视频创建请求。

示意性的，音频应用为音乐播放应用。

终端 110 与服务器 120 之间通过通信网络相连。可选的，通信网络是有线网络或无线网络。

服务器 120 是一台服务器，或者由若干台服务器组成的服务器集群，或者是一个虚拟化平台，或者是一个云计算服务中心。

可选的，服务器 120 包括提供音视频处理服务的后台服务器。可选的，服务器为音频应用对应的后台服务器。

服务器 120 与终端 110 之间通过通信网络相连。可选的，通信网络是有线网络或无线网络。

服务器 120 用于接收终端 110 发送的视频创建请求，从该视频创建请求中获取目标视频，并存储该目标视频。

可选的，该终端 110 是具有视频播放功能的电子设备。该终端 110，还用于向服务器 120 发送目标音频的音频播放请求，根据服务器 120 反馈的画面播放指令，通过音频应用在播放目标音频时播放目标视频的画面信息。

需要说明的是，上述的上传视频的终端与播放音频的终端可以是同一个终端，也可以是不同的两个终端，本实施例对此不加以限定。为了方便说明，仅以上传视频的终端与播放音频的终端为同一个终端，即本申请实施例中涉及的终端是同时具有视频上传功能和视频播放功能的终端为例进行说明。

请参考图 2, 其示出了本申请一个示例性实施例提供的音视频处理方法的流程图。本实施例以该音视频处理方法应用于图 1 所示出的终端中来举例说明。该音视频处理方法包括:

步骤 201, 显示目标音频的视频创建界面, 视频创建界面中包括目标音频的 n 个音频片段和 n 个音频片段各自对应的视频录制入口, 其中, $n \geq 2$ 。

当音频应用处于前台运行状态时, 终端通过该音频应用显示目标音频的视频创建界面。

可选的, 目标音频的视频创建界面是用于为目标音频中的至少一个音频片段创建对应的目标视频的界面。

视频创建界面中包括目标音频的 n 个音频片段和该 n 个音频片段各自对应的视频录制入口, 可选地, 该 n 个音频片段可以是组成完整目标音频的 n 个音频片段, 也可以是组成部分目标音频的 n 个音频片段。每个音频片段对应的视频录制入口可以直接显示在视频创建界面中, 也可以在终端接收到视频创建界面中作用于音频片段的选中信号时显示在该音频片段上。

可选的, 视频录制入口为用于触发获取目标音频片段对应的目标视频的可操作控件。示意性的, 视频录制入口的类型包括按钮、可操控的条目、滑块中的至少一种。本实施例对视频录制入口的位置和类型均不加以限定。

可选的, 终端通过音频应用播放目标音频。该目标音频中包括多个音频片段。多个音频片段中存在至少两个音频片段各自对应的播放时长是不相同的, 或者各个音频片段对应的播放时长均相同。比如, 多个音频片段中的每个音频片段对应的播放时长均为 30 秒。

可选的, 终端显示目标音频的视频创建界面之前, 服务器预先对目标音频进行分片处理。为了保证分片处理后得到的多个音频片段中歌词的完整性, 服务器根据目标音频的节拍特征对目标音频进行分片处理, 得到多个音频片段。

其中, 节拍特征用于指示视频对应的每分钟节拍数 (Beat Per Minute, BPM)。

步骤 202, 在视频创建界面中接收作用于目标视频录制入口的触发信号, 该目标视频录制入口为与目标音频片段对应的视频录制入口。

在一种可能的实现方式中, 当终端接收到视频创建界面中作用于音频片段的选中信号时, 将该音频片段确定为目标音频片段, 当终端接收到作用于目标音频片段上的目标视频录制入口的触发信号时, 执行下述步骤 203。

触发信号是用于触发获取目标音频片段对应的目标视频的用户操作。示意性的，触发信号包括点击操作信号、滑动操作信号、按压操作信号、长按操作信号中的任意一种或多种的组合。

可选的，触发信号还包括其它可能的实现方式。比如，触发信号以语音形式实现，或者触发信号以摇一摇操作的形式实现，摇一摇操作是指摇晃终端的操作。

步骤 203，根据触发信号获取目标音频片段对应的目标视频，目标视频为视频时长小于时长阈值的视频片段。

可选的，终端获取目标音频片段对应的目标视频，可以实现为终端获取至少两个目标音频片段各自对应的目标视频。

可选的，目标音频片段为目标音频的多个音频片段中的一个或者多个音频片段，一个目标音频片段对应的目标视频包括一个目标音频片段对应的一个或者多个目标视频。本申请实施例对目标音频片段和目标视频的数量均不加以限定。

其中，目标视频为视频时长小于时长阈值的视频片段。可选的，目标视频为终端录制或者上传的视频。比如，时长阈值为 15 秒、20 秒或者 1 分钟。视频为原创的。

可选的，目标视频为短视频，短视频是一种互联网内容的传播方式。目标视频包括画面信息和音频信息。目标视频是画面信息和/或音频信息为原创内容且视频时长小于时长阈值的视频。示意性的，目标视频是画面信息由用户录制的且音频信息的时长小于时长阈值的视频。示意性的，目标视频是画面信息和音频信息均由用户录制的且视频时长小于时长阈值的视频。

步骤 204，向服务器发送携带有目标视频的视频创建请求，视频创建请求用于指示在播放目标音频时播放目标视频的画面信息。

终端根据目标音频片段对应的目标视频，生成视频创建请求，该视频创建请求携带有目标视频。终端将该视频创建请求发送至服务器。对应的，服务器接收视频创建请求，从视频创建请求中获取目标视频，并对目标视频进行存储。

可选的，当终端通过音频应用播放该目标音频时，从服务器中获取该目标音频对应的目标视频的画面信息，在播放目标音频时播放目标视频的画面信息，此时目标视频的音频信息是不播放的。

综上所述，本实施例通过显示目标音频的视频创建界面，在视频创建界面

中接收作用于目标音频片段对应的视频录制入口的触发信号，根据触发信号获取目标音频片段对应的目标视频，向服务器发送携带有目标视频的视频创建请求，使得终端可以针对一个音频中的一个或者多个音频片段上传相应的视频，进而使得后续在播放该音频的过程中能够同步播放音频片段对应的视频的画面信息，避免了相关技术中音频对应的 MV 是固定不可变的导致音视频播放效果差的问题，提高了音视频的播放效果。

请参考图 3，其示出了本申请另一个示例性实施例提供的音视频处理方法的流程图。本实施例以该音视频处理方法应用于图 1 所示出的终端中来举例说明。该音视频处理方法包括：

步骤 301，终端显示目标音频的音频关联界面，音频关联界面中包括视频创建入口。

可选的，音频关联界面为与目标音频的播放过程相关联的界面。示意性的，目标音频的音频关联界面为目标音频的音频播放界面。

视频创建入口是音频应用中提供的入口。视频创建入口是用于触发显示视频创建界面的可操作控件。示意性的，视频创建入口的类型包括按钮、可操控的条目、滑块中的至少一种。本实施例对视频创建入口的位置和类型均不加以限定。

步骤 302，当终端接收到作用于视频创建入口的触发信号时，显示目标音频的视频创建界面。

当终端接收到音频关联界面上的视频创建入口对应的触发信号时，显示目标音频的视频创建界面。

触发信号是用于触发显示目标音频的视频创建界面的用户操作。示意性的，触发信号包括点击操作信号、滑动操作信号、按压操作信号、长按操作信号中的任意一种或多种的组合。

目标音频的视频创建界面中包括目标音频的多个音频片段和多个音频片段各自对应的视频录制入口。

步骤 303，终端在视频创建界面中接收作用于目标视频录制入口的触发信号，该目标视频录制入口为与目标音频片段对应的视频录制入口。

需要说明的是，终端在视频创建界面中接收作用于目标视频录制入口的触发信号的过程，请参考上述实施例中的相关细节，在此不再赘述。

步骤 304, 终端根据触发信号获取目标音频片段对应的目标视频。

其中, 目标视频为视频时长小于时长阈值的视频片段。

终端根据触发信号获取目标音频片段对应的目标视频, 包括但不限于以下几种可能的实现方式:

在一种可能的实现方式中, 终端根据触发信号开启视频录制功能, 进行视频录制得到目标视频。

可选的, 终端在开启视频录制功能之后, 按照预设采样频率采集图像, 当终端检测到停止控件对应的停止操作信号时, 结束视频录制功能, 同时对采集到的各帧图像进行编码, 合成得到目标视频。

其中, 停止控件可以实现为在视频录制过程中在终端的屏幕上显示的控件, 也可以实现为终端上的物理按键。停止控件是用于结束视频录制功能的可操作控件。示意性的, 当该停止控件实现为屏幕上显示的控件时, 该停止控件的类型包括按钮、可操控的条目、滑块中的至少一种。本实施例对停止控件的位置和类型均不加以限定。

停止操作信号是用于触发结束停止控件对应的视频录制功能的用户操作信号。示意性的, 停止操作信号包括点击操作信号、滑动操作信号、按压操作信号、长按操作信号中的任意一种或多种的组合。

可选的, 停止操作还包括其它可能的实现方式。比如, 停止操作以语音形式实现, 或者停止操作以摇一摇操作的形式实现, 摇一摇操作是指摇晃终端的操作。

在另一种可能的实现方式中, 终端根据触发信号获取终端中存储的目标视频。

可选的, 终端中存储有已录制或者已下载的目标视频。终端根据触发信号获取终端中存储的目标视频。

在另一种可能的实现方式中, 终端根据触发信号向服务器发送视频获取请求, 视频获取请求中携带有目标音频片段的片段标识, 视频获取请求用于指示服务器反馈片段标识对应的至少一个视频; 从接收到的服务器反馈的至少一个视频中确定目标视频。

终端根据触发信号向服务器发送视频获取请求, 对应的, 服务器接收该视频获取请求, 从该视频获取请求中获取目标音频片段的片段标识, 根据第一对应关系获取该片段标识对应的多个视频, 向终端反馈该片段标识对应的多个视

频。对应的，终端接收服务器反馈的多个视频，从多个视频中确定目标视频。

服务器中存储有目标音频的多个音频片段的片段标识与多个视频之间的对应关系。多个音频片段中存在至少一个音频片段的片段标识对应多个视频。

终端从多个视频中确定目标视频，包括：终端从多个视频中随机确定目标视频；终端根据视频信息从多个视频中确定目标视频。其中，视频信息包括视频的点赞数、视频的更新时间和视频的创建时间中的至少一种。

需要说明的是，目标视频是在终端接收到作用于目标音频片段对应的视频录制入口的触发信号后进行录制并上传的视频，或者是在终端接收到作用于目标音频片段对应的视频录制入口的触发信号时上传的已存储的视频，或者是在终端接收到作用于目标音频片段对应的视频录制入口的触发信号时从服务器反馈的多个视频中确定的视频。本实施例对目标视频的获取方式不加以限定。

步骤 305，终端向服务器发送携带有目标视频的视频创建请求。

视频创建请求用于指示在播放目标音频时播放目标视频的画面信息。

可选的，该目标音频片段还可以实现为至少两个目标音频片段，该至少两个目标视频片段各自对应目标视频。其中，至少两个目标音频片段为目标音频中连续的至少两个音频片段。

可选的，终端向服务器发送携带有目标视频的视频创建请求包括：终端将至少两个目标视频各自对应的画面信息进行拼接得到第一视频拼接画面，向服务器发送携带有第一视频拼接画面的第一视频创建请求，第一视频创建请求用于指示在播放目标音频时播放第一视频拼接画面。

可选的，终端将至少两个目标视频各自对应的画面信息进行拼接得到第一视频拼接画面之前，还包括：将播放时长大于对应的目标音频片段的播放时长的目标视频进行剪裁。

当存在目标视频的第一播放时长大于对应的目标音频片段的播放时长时，终端将具有第一播放时长的目标视频剪裁至具有第二播放时长的目标视频。

示意性的，当存在目标视频的第一播放时长大于对应的目标音频片段的播放时长时，服务器确定目标视频的第一播放时长与对应的目标音频片段的播放时长的差值绝对值，服务器将目标视频的视频头部或者视频尾部剪裁出具有差值绝对值的播放时长的片段，得到具有第二播放时长的目标视频。

比如，目标视频的播放时长为 40 秒，对应的音频片段的播放时长为 30 秒，服务器确定差值绝对值为 10，将目标视频的播放时长剪裁出 10 秒，得到具有 3

0 秒播放时长的目标视频。

步骤 306, 服务器接收终端发送的视频创建请求, 将视频创建请求中的目标视频进行存储。

可选的, 服务器接收终端发送的第一视频创建请求, 从第一视频创建请求中获取第一视频拼接画面, 并存储该第一视频拼接画面。

可选的, 服务器中存储目标音频的音频标识与该第一视频拼接画面之间的第四对应关系, 音频标识用于在多个音频中唯一标识该目标音频。

在一个示意性的例子中, 如图 4 所示, 终端显示目标音频的视频创建界面 41, 视频创建界面 41 中包括目标音频“123 我爱你”的多个音频片段。当终端在视频创建界面 41 中接收作用于目标音频片段 42 对应的视频录制入口 43 的点击信号时, 终端开启视频录制功能, 进行视频录制得到目标音频片段 42 对应的目标视频, 终端向服务器发送携带有该目标视频的视频创建请求。

可选的, 如图 5 所示, 步骤 306 之后包括但不限于以下几个步骤:

步骤 501, 终端向服务器发送目标音频的音频播放请求。

其中, 音频播放请求用于指示服务器获取 n 个音频片段各自对应的第一候选视频并反馈画面播放指令, 该第一候选视频包括目标视频; 向终端发送画面播放指令。

可选的, 音频播放请求携带有目标音频的音频标识。目标音频的音频标识用于唯一指示该目标音频。

步骤 502, 服务器接收终端发送的音频播放请求。

对应的, 服务器接收终端发送的音频播放请求, 从音频播放请求中获取目标音频的音频标识。

步骤 503, 服务器获取与目标音频匹配的多个候选视频, 多个候选视频中包括目标视频。

可选的, 服务器获取与目标音频匹配的多个候选视频包括但不限于以下几种可能的实现方式:

在一种可能的实现方式中, 服务器获取目标音频的音频标签, 音频标签用于指示目标音频的类型; 根据第二对应关系获取与音频标签对应的目标视频分类组, 第二对应关系包括音频标签与视频分类组之间的对应关系; 从目标视频分类组中获取多个候选视频。

需要说明的是，本申请实施例中涉及的视频（包括目标视频、候选视频）均为短视频，短视频为视频时长小于时长阈值的视频片段。

目标音频的音频标签可以是服务器根据目标音频的音频内容确定的，也可以是服务器接收到的终端发送的音频播放请求中携带的。本实施例对此不加以限定。

音频标签用于唯一指示音频的音频类型。在一种可能的实现方式中，按照音频素材的不同对音频进行分类，比如，音频类型包括风景类型、舞蹈类型和演唱类型中的至少一种。在另一种可能的实现方式中，按照音频节奏的不同对音频进行分类，比如，音频类型包括快节奏类型和慢节奏类型。本实施例对音频的分类方式不加以限定。

需要说明的是，本申请实施例对音频的分类方式可以与对视频的分类方式是相同的或者是不同的。为了方便介绍，下面仅以对音频的分类方式和对视频的分类方式是相同的，即音频类型和视频类型均包括风景类型、舞蹈类型和演唱类型这三种类型为例进行说明。

示意性的，音频标签与视频分类组之间的对应关系如表一所示。音频标签包括风景标签、舞蹈标签和演唱标签这三种，与风景标签匹配的目标视频分类组为第一视频分类组，该第一视频分类组中包括多个具有风景标识的视频；与舞蹈标签匹配的目标视频分类组为第二视频分类组，该第二视频分类组中包括多个具有舞蹈标识的视频；与演唱标签匹配的目标视频分类组为第三视频分类组，该第三视频分类组中包括多个具有演唱标识的视频。

表一

音频标签	视频分类组
风景标识	第一视频分类组
舞蹈标识	第二视频分类组
演唱标识	第三视频分类组

可选的，服务器从目标视频分类组中随机获取多个候选视频，多个候选视频的第三播放时长之和小于或等于目标音频的播放时长。

在另一种可能的实现方式中，服务器获取与目标音频匹配的多个候选视频，包括：服务器获取 n 个音频片段各自对应的第一候选视频，第一候选视频包括所述目标视频。

可选的，服务器根据第三对应关系从视频库中获取多个音频片段各自对应

的第一候选视频，第三对应关系包括音频片段与第一候选视频之间的对应关系。

可选的，服务器中存储有目标音频的多个音频片段各自对应的候选视频集，每个音频片段对应的候选视频集中包括至少一个候选视频。音频片段对应的第一候选视频为该音频片段对应的候选视频集中的任意一个。

示意性的，目标音频为歌曲“AA”，歌曲“AA”包括5个音频片段，每个音频片段的播放时长为30秒，5个音频片段与多个候选视频集之间的对应关系如表二所示。片段1对应的第一候选视频集，第一候选视频集包括“视频S001、视频S025、视频S067、视频S091、视频S101、视频S134、视频S175”；片段2对应的第二候选视频集，第二候选视频集包括“视频S010、视频S106”；片段3对应的第三候选视频集，第三候选视频集包括“视频S003、视频S012、视频S050、视频S079、视频S111”；片段4对应的第四候选视频集，第四候选视频集包括“视频S007、视频S088”；片段5对应的第五候选视频集，第五候选视频集包括“视频S008、视频S053、视频S099、视频S190、视频S351”。

表二

音频片段	候选视频集
片段1	视频S001、视频S025、视频S067、视频S091、视频S101、视频S134、视频S175
片段2	视频S010、视频S106
片段3	视频S003、视频S012、视频S050、视频S079、视频S111
片段4	视频S007、视频S088
片段5	视频S008、视频S053、视频S099、视频S190、视频S351

在另一种可能的实现方式中，服务器获取目标音频的音频标识，根据第四对应关系获取与音频标识对应的第一视频拼接画面。

步骤504，服务器向终端发送画面播放指令。

画面播放指令用于指示音频应用在播放目标音频时播放多个候选视频各自对应的画面信息。

可选的，服务器向终端发送画面播放指令，包括但不限于以下几种可能的发送方式：

在一种可能的实现方式中，当服务器获取到的是多个候选视频各自对应的画面信息，且多个候选视频的第三播放时长之和小于或等于目标音频的播放时长时，服务器向终端发送第一画面播放指令，第一画面播放指令用于指示在播放目标音频的过程中顺序播放多个候选视频各自对应的画面信息。

在另一种可能的实现方式中，当服务器获取到的是多个候选视频各自对应的画面信息，且多个候选视频的第三播放时长之和大于目标音频的播放时长时，服务器将具有第三播放时长之和的多个候选视频剪裁至具有第四播放时长之和的多个候选视频，向终端发送第二画面播放指令。

当多个候选视频的第三播放时长之和大于目标音频的播放时长时，服务器先对多个候选视频进行剪裁，即将具有第三播放时长之和的多个候选视频剪裁至具有第四播放时长之和的多个候选视频。其中，第四播放时长之和小于或者等于目标音频的播放时长。

可选的，服务器确定多个候选视频的第三播放时长之和与目标音频的播放时长的差值绝对值，服务器将多个候选视频中的一个或者多个候选视频进行剪裁，裁出具有差值绝对值的播放时长的片段，得到具有第四播放时长之和的多个候选视频。

可选的，第二画面播放指令用于指示在播放目标音频的过程中顺序播放具有第四播放时长之和的多个候选视频各自对应的画面信息。

在另一种可能的实现方式中，当服务器获取到的是多个候选视频各自对应的画面信息，且多个候选视频的第三播放时长之和大于目标音频的播放时长时，服务器向终端发送第三画面播放指令，第三画面播放指令用于指示在播放目标音频的过程中顺序播放多个候选视频各自对应的画面信息，当目标音频播放结束时停止对画面信息的播放。

由于多个候选视频的第三播放时长之和大于目标音频的播放时长，即当目标音频播放结束时画面信息仍继续播放，因此服务器可以指示当目标音频播放结束时停止对画面信息的播放。

在另一种可能的实现方式中，当服务器获取到的是多个音频片段各自对应的第一候选视频时，服务器向终端发送画面播放指令，包括：向终端发送第四画面播放指令，第四画面播放指令用于指示音频应用在播放目标音频时播放多个第一候选视频各自对应的画面信息。

可选的，服务器向终端发送第四画面播放指令，第四画面播放指令还用于

指示在播放目标音频时顺序播放多个第一候选视频各自对应的画面信息。

在另一种可能的实现方式中，当服务器获取到的是拼接好的第一视频拼接画面时，服务器向终端发送第五画面播放指令，第五画面播放指令用于指示在播放目标音频时播放第一视频拼接画面。

步骤 505，终端根据接收到的画面播放指令，在播放目标音频的音频片段时播放 n 个音频片段各自对应的第一候选视频的画面信息。

可选的，当终端接收到第一画面播放指令时，通过音频应用在播放目标音频的过程中顺序播放 n 个候选视频各自对应的画面信息。

可选的，当终端接收到第二画面播放指令时，通过音频应用在播放目标音频的过程中顺序播放具有第四播放时长之和的多个候选视频各自对应的画面信息。

可选的，当终端接收到第三画面播放指令时，通过音频应用在播放目标音频的过程中顺序播放多个候选视频各自对应的画面信息，当目标音频播放结束时停止对画面信息的播放。

可选的，当终端接收到第四画面播放指令时，通过音频应用在播放目标音频时播放多个第一候选视频各自对应的画面信息。

可选的，当终端接收到第五画面播放指令时，通过音频应用在播放目标音频时播放第一视频拼接画面。

由于终端播放目标视频的过程中可能存在某个音频片段对应的画面信息不是用户所喜欢的，因此用户可以通过画面切换操作切换播放其他的画面信息。在一种可能的实现方式中，当终端接收到作用于第一音频片段的画面切换操作信号时，获取目标音频当前的第一播放时间点；向服务器发送画面切换请求，画面切换请求携带有目标音频的音频标识，画面切换请求用于指示服务器获取与音频标识对应的第二视频拼接画面；向终端发送第一画面切换指令；终端根据接收到的第一画面切换指令，在目标音频的第一播放时间点上开始切换播放第二视频拼接画面。

第一播放时间点为终端接收到作用于第一音频片段的画面切换操作信号时的目标音频的播放时间点。

可选的，终端向服务器发送画面切换请求，对应的，服务器接收该画面切换请求，从画面切换请求中获取音频标识，获取与音频标识对应的第二视频拼接画面。服务器向终端反馈第一画面切换指令。终端接收服务器反馈的第一画

面切换指令，根据该第一画面切换指令在目标音频的第一播放时间点上开始切换播放第二视频拼接画面。

其中，第二视频拼接画面与第一视频拼接画面不同。

可选的，终端获取与音频标识对应的第二视频拼接画面，包括：终端获取与音频标识对应的候选视频拼接画面集，将候选视频拼接画面集中除了第一视频拼接画面以外的其它任意一个视频拼接画面确定为第二视频拼接画面。

示意性的，终端在进行画面切换时，无需从头开始播放目标音频，而是继续播放该目标音频，在目标音频的第一播放时间点上将第一视频拼接画面切换播放为第二视频拼接画面。

在一种可能的实现方式中，画面切换请求中还携带有第一播放时间点，画面切换请求还用于指示服务器当不存在与音频标识对应的第二视频拼接画面时，获取与第一播放时间点所在的第一音频片段对应的第二候选视频；向终端发送第二画面切换指令。

可选的，当服务器中不存在与音频标识对应的第二视频拼接画面时，服务器获取与第一播放时间点所在的第一音频片段对应的第二候选视频，向终端发送第二画面切换指令。其中，第二画面切换指令用于指示在播放第一音频片段的过程中在目标音频的第一播放时间点上开始切换播放第二候选视频的画面信息。

示意性的，服务器获取与第一播放时间点所在的第一音频片段对应的第二候选视频，包括：获取与第一目标音频片段对应的候选视频集，将候选视频集中除了第一候选视频以外的其它任意一个候选视频确定为第二候选视频。

可选的，终端根据接收到的第二画面切换指令，在播放第一音频片段的过程中在目标音频的第一播放时间点上开始切换播放第二候选视频的画面信息。

可选的，音频片段的播放时长与对应的视频的播放时长是相同的，音频片段与对应的视频是同步播放的，即音频片段的播放时间点与视频的播放时间点存在对应关系。因此，终端接收服务器发送的画面切换指令时，终端确定目标音频的目标播放时间点对应的第二候选视频的目标播放时间点，终端通过音频应用在目标播放时间点上继续播放第一目标音频片段的同时，切换播放第二候选视频的画面信息，切换播放的第二候选视频的画面信息是从第二候选视频的目标播放时间点开始的画面信息。

由于当终端在播放目标音频的前奏内容、间奏内容和尾奏内容中的至少一

种时，无需播放相应的目标视频的画面信息。因此，在一种可能的实现方式中，终端在播放目标音频的指定时间内容时，将第一候选视频的画面信息切换显示为所述第一候选视频的视频制作信息，该第一候选视频为指定时间内容所在的音频片段对应的候选视频，指定时间内容包括前奏内容、间奏内容和尾奏内容中的至少一种，视频制作信息包括第一候选视频对应的录制者信息、上传者信息和拼接者信息中的至少一种。

示意性的，录制者信息为录制该第一候选视频的帐号标识，上传者信息为将该第一候选视频上传至服务器的帐号标识。当第一候选视频的画面信息为已拼接好的视频拼接画面的部分时，拼接者信息为对将多个画面信息拼接为视频拼接画面的帐号标识。

可选的，音频片段的播放时长与对应的候选视频的播放时长是相同的，音频片段与对应的候选视频是同步播放的，即音频片段的播放时间点与候选视频的播放时间点存在对应关系。因此，终端接收服务器发送的画面切换指令时，终端确定目标音频的目标播放时间点对应的第二候选视频的目标播放时间点，终端通过音频应用在目标播放时间点上继续播放第一目标音频片段的同时，切换播放第二候选视频的画面信息，切换播放的第二候选视频的画面信息是从第二候选视频的目标播放时间点开始的画面信息。

综上所述，本申请实施例还通过终端将多个目标视频各自对应的画面信息进行拼接得到第一视频拼接画面，向服务器发送携带有第一视频拼接画面的第一视频创建请求，第一视频创建请求用于指示在播放目标音频时播放第一视频拼接画面；由于终端预先为目标音频创建好拼接完成的视频拼接画面，使得后续播放该目标音频时可以直接播放创建好的视频拼接画面，进而使得在播放目标音频的过程中画面显示更加流畅，进一步提高了目标音频的播放效果。

本申请实施例还通过服务器接收终端发送的音频播放请求，音频播放请求用于请求开始播放目标音频；获取与目标音频匹配的多个候选视频，多个候选视频中包括目标视频；向终端发送画面播放指令，使得终端根据接收到的画面播放指令，通过音频应用播放目标音频时播放多个候选视频各自对应的画面信息，达到了目标音频与上传至服务器中的多个候选视频的画面信息进行同步播放的技术效果。

本申请实施例还通过当接收到作用于第一音频片段的画面切换操作信号时，获取目标音频当前的第一播放时间点；向服务器发送画面切换请求，画面

切换请求携带有目标音频的音频标识，画面切换请求用于指示服务器获取与音频标识对应的第二视频拼接画面；向终端发送第一画面切换指令；根据接收到的第一画面切换指令，在目标音频的第一播放时间点上开始切换播放第二视频拼接画面；避免了画面切换播放时需要从头开始播放的情况，使得第二终端能够在目标播放时间点上继续播放切换的第二视频拼接画面，提高了目标音频的播放效果。

下述为本申请装置实施例，可以用于执行本申请方法实施例。对于本申请装置实施例中未披露的细节，请参照本申请方法实施例。

请参考图 6，其示出了本申请一个示例性实施例提供的音视频处理装置的结构示意图。该音视频处理装置可以通过专用硬件电路，或者，软硬件的结合实现成为图 1 中的终端的全部或一部分，该音视频处理装置包括：显示模块 610、接收模块 620、获取模块 630 和发送模块 640。

显示模块 610，用于显示目标音频的视频创建界面，所述视频创建界面中包括所述目标音频的 n 个音频片段和所述 n 个音频片段各自对应的视频录制入口，其中， $n \geq 2$ ；

接收模块 620，用于在所述视频创建界面中接收作用于目标视频录制入口的触发信号，所述目标视频录制入口为与目标音频片段对应的视频录制入口；

获取模块 630，用于根据触发信号获取目标音频片段对应的目标视频，目标视频为视频时长小于时长阈值的视频片段；

发送模块 640，用于向服务器发送携带有目标视频的视频创建请求，视频创建请求用于指示在播放目标音频时播放目标视频的画面信息。

可选的，获取模块 630，还用于根据触发信号开启视频录制功能，进行视频录制得到目标视频；或者，

根据触发信号获取终端中存储的目标视频；或者，

根据所述触发信号向所述服务器发送视频获取请求，所述视频获取请求中携带有所述目标音频片段的片段标识，所述视频获取请求用于指示所述服务器反馈所述片段标识对应的至少一个视频；从接收到的所述服务器反馈的所述至少一个视频中确定所述目标视频。

可选的，显示模块 610，还用于显示目标音频的音频关联界面，音频关联界

面中包括视频创建入口；

当接收到作用于视频创建入口的触发信号时，显示目标音频的视频创建界面。

可选的，所述目标音频片段包括至少两个所述目标音频片段，所述至少两个目标音频片段各自对应有所述目标视频；

如图 7 所示，该装置还包括：

拼接模块 650，用于将多个目标视频各自对应的画面信息进行拼接得到第一视频拼接画面；

发送模块 640，还用于向服务器发送携带有第一视频拼接画面的第一视频创建请求，第一视频创建请求用于指示在播放目标音频时播放第一视频拼接画面。

可选的，该装置还包括：

播放模块 660，用于向所述服务器发送所述目标音频的音频播放请求，所述音频播放请求用于指示所述服务器获取所述 n 个音频片段各自对应的第一候选视频并反馈画面播放指令，所述第一候选视频包括所述目标视频；根据接收到的所述画面播放指令，在播放所述目标音频的音频片段时播放所述 n 个音频片段各自对应的所述第一候选视频的画面信息。

可选的，装置还包括：

第一画面切换模块 670，用于当接收到作用于第一音频片段的画面切换操作信号时，获取目标音频当前的第一播放时间点；

向服务器发送画面切换请求，画面切换请求携带有目标音频的音频标识，画面切换请求用于指示服务器获取与音频标识对应的第二视频拼接画面；向终端发送第一画面切换指令；

根据接收到的第一画面切换指令，在目标音频的第一播放时间点上开始切换播放第二视频拼接画面。

可选的，画面切换请求中还携带有第一播放时间点，画面切换请求还用于指示服务器当不存在与音频标识对应的第二视频拼接画面时，获取与第一播放时间点所在的第一音频片段对应的第二候选视频；向终端发送第二画面切换指令；

该装置还包括：

第二画面切换模块 680，用于根据接收到的第二画面切换指令，在播放第一音频片段的过程中在目标音频的第一播放时间点上开始切换播放第二候选视频

的画面信息。

可选的，该装置还包括：

切换显示模块 690，用于在播放所述目标音频的指定时间内容时，将第一候选视频的画面信息切换显示为所述第一候选视频的视频制作信息，所述第一候选视频为所述指定时间内容所在的音频片段对应的候选视频；所述指定时间内容包括前奏内容、间奏内容和尾奏内容中的至少一种，所述视频制作信息包括所述第一候选视频对应的录制者信息、上传者信息和拼接者信息中的至少一种。

相关细节可结合参考图 2 至图 5 所示的方法实施例。其中，显示模块 610 还用于实现上述方法实施例中其他任意隐含或公开的与显示步骤相关的功能；接收模块 620 还用于实现上述方法实施例中其他任意隐含或公开的与接收步骤相关的功能；获取模块 630 还用于实现上述方法实施例中其他任意隐含或公开的与获取步骤相关的功能；发送模块 640 还用于实现上述方法实施例中其他任意隐含或公开的与发送步骤相关的功能。

需要说明的是，上述实施例提供的装置，在实现其功能时，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将设备的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。另外，上述实施例提供的装置与方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

图 8 示出了本申请一个示例性实施例提供的终端 800 的结构框图。该终端 800 可以是：智能手机、平板电脑、MP3 播放器（Moving Picture Experts Group Audio Layer III，动态影像专家压缩标准音频层面 3）、MP4（Moving Picture Experts Group Audio Layer IV，动态影像专家压缩标准音频层面 4）播放器、笔记本电脑或台式电脑。终端 800 还可能被称为用户设备、便携式终端、膝上型终端、台式终端等其他名称。

通常，终端 800 包括有：处理器 801 和存储器 802。

处理器 801 可以包括一个或多个处理核心，比如 4 核心处理器、8 核心处理器等。处理器 801 可以采用 DSP(Digital Signal Processing, 数字信号处理)、FPGA (Field-Programmable Gate Array, 现场可编程门阵列)、PLA (Programmable Logic Array, 可编程逻辑阵列) 中的至少一种硬件形式来实现。处理器 801 也可以包括主处理器和协处理器，主处理器是用于对在唤醒状态下的数据进行处理

的处理器，也称 CPU（Central Processing Unit，中央处理器）；协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中，处理器 801 可以在集成有 GPU（Graphics Processing Unit，图像处理器），GPU 用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中，处理器 801 还可以包括 AI（Artificial Intelligence，人工智能）处理器，该 AI 处理器用于处理有关机器学习的计算操作。

存储器 802 可以包括一个或多个计算机可读存储介质，该计算机可读存储介质可以是非暂态的。存储器 802 还可包括高速随机存取存储器，以及非易失性存储器，比如一个或多个磁盘存储设备、闪存存储设备。在一些实施例中，存储器 802 中的非暂态的计算机可读存储介质用于存储至少一个指令，该至少一个指令用于被处理器 801 所执行以实现本申请中方法实施例提供的音视频处理方法。

在一些实施例中，终端 800 还可选包括有：外围设备接口 803 和至少一个外围设备。处理器 801、存储器 802 和外围设备接口 803 之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外围设备接口 803 相连。具体地，外围设备包括：射频电路 804、触摸显示屏 805、摄像头 806、音频电路 807、定位组件 808 和电源 809 中的至少一种。

外围设备接口 803 可被用于将 I/O（Input /Output，输入/输出）相关的至少一个外围设备连接到处理器 801 和存储器 802。在一些实施例中，处理器 801、存储器 802 和外围设备接口 803 被集成在同一芯片或电路板上；在一些其他实施例中，处理器 801、存储器 802 和外围设备接口 803 中的任意一个或两个可以在单独的芯片或电路板上实现，本实施例对此不加以限定。

射频电路 804 用于接收和发射 RF（Radio Frequency，射频）信号，也称电磁信号。射频电路 804 通过电磁信号与通信网络以及其他通信设备进行通信。射频电路 804 将电信号转换为电磁信号进行发送，或者，将接收到的电磁信号转换为电信号。可选地，射频电路 804 包括：天线系统、RF 收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频电路 804 可以通过至少一种无线通信协议来与其它终端进行通信。该无线通信协议包括但不限于：万维网、城域网、内联网、各代移动通信网络（2G、3G、4G 及 5G）、无线局域网和/或 WiFi（Wireless Fidelity，无线保真）网络。在一些实施例中，射频电路 804 还可以包括 NFC（Near Field Communication，

近距离无线通信)有关的电路,本申请对此不加以限定。

显示屏 805 用于显示 UI (User Interface, 用户界面)。该 UI 可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏 805 是触摸显示屏时,显示屏 805 还具有采集在显示屏 805 的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器 801 进行处理。此时,显示屏 805 还可以用于提供虚拟按钮和/或虚拟键盘,也称软按钮和/或软键盘。在一些实施例中,显示屏 805 可以为一个,设置终端 800 的前面板;在另一些实施例中,显示屏 805 可以为至少两个,分别设置在终端 800 的不同表面或呈折叠设计;在再一些实施例中,显示屏 805 可以是柔性显示屏,设置在终端 800 的弯曲表面上或折叠面上。甚至,显示屏 805 还可以设置成非矩形的不规则图形,也即异形屏。显示屏 805 可以采用 LCD(Liquid Crystal Display, 液晶显示屏)、OLED(Organic Light-Emitting Diode, 有机发光二极管)等材质制备。

摄像头组件 806 用于采集图像或视频。可选地,摄像头组件 806 包括前置摄像头和后置摄像头。通常,前置摄像头设置在终端的前面板,后置摄像头设置在终端的背面。在一些实施例中,后置摄像头为至少两个,分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种,以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及 VR (Virtual Reality, 虚拟现实)拍摄功能或者其它融合拍摄功能。在一些实施例中,摄像头组件 806 还可以包括闪光灯。闪光灯可以是单色温闪光灯,也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合,可以用于不同色温下的光线补偿。

音频电路 807 可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波,并将声波转换为电信号输入至处理器 801 进行处理,或者输入至射频电路 804 以实现语音通信。出于立体声采集或降噪的目的,麦克风可以为多个,分别设置在终端 800 的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器 801 或射频电路 804 的电信号转换为声波。扬声器可以是传统的薄膜扬声器,也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时,不仅可以将电信号转换为人类可听见的声波,也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中,音频电路 807 还可以包括耳机插孔。

定位组件 808 用于定位终端 800 的当前地理位置,以实现导航或 LBS

(Location Based Service, 基于位置的服务)。定位组件 808 可以是基于美国的 GPS (Global Positioning System, 全球定位系统)、中国的北斗系统或俄罗斯的伽利略系统的定位组件。

电源 809 用于为终端 800 中的各个组件进行供电。电源 809 可以是交流电、直流电、一次性电池或可充电电池。当电源 809 包括可充电电池时, 该可充电电池可以是有线充电电池或无线充电电池。有线充电电池是通过有线线路充电的电池, 无线充电电池是通过无线线圈充电的电池。该可充电电池还可以用于支持快充技术。

在一些实施例中, 终端 800 还包括有一个或多个传感器 810。该一个或多个传感器 810 包括但不限于: 加速度传感器 811、陀螺仪传感器 812、压力传感器 813、指纹传感器 814、光学传感器 815 以及接近传感器 816。

加速度传感器 811 可以检测以终端 800 建立的坐标系的三个坐标轴上的加速度大小。比如, 加速度传感器 811 可以用于检测重力加速度在三个坐标轴上的分量。处理器 801 可以根据加速度传感器 811 采集的重力加速度信号, 控制触摸显示屏 805 以横向视图或纵向视图进行用户界面的显示。加速度传感器 811 还可以用于游戏或者用户的运动数据的采集。

陀螺仪传感器 812 可以检测终端 800 的机体方向及转动角度, 陀螺仪传感器 812 可以与加速度传感器 811 协同采集用户对终端 800 的 3D 动作。处理器 801 根据陀螺仪传感器 812 采集的数据, 可以实现如下功能: 动作感应 (比如根据用户的倾斜操作来改变 UI)、拍摄时的图像稳定、游戏控制以及惯性导航。

压力传感器 813 可以设置在终端 800 的侧边框和/或触摸显示屏 805 的下层。当压力传感器 813 设置在终端 800 的侧边框时, 可以检测用户对终端 800 的握持信号, 由处理器 801 根据压力传感器 813 采集的握持信号进行左右手识别或快捷操作。当压力传感器 813 设置在触摸显示屏 805 的下层时, 由处理器 801 根据用户对触摸显示屏 805 的压力操作, 实现对 UI 界面上的可操作性控件进行控制。可操作性控件包括按钮控件、滚动条控件、图标控件、菜单控件中的至少一种。

指纹传感器 814 用于采集用户的指纹, 由处理器 801 根据指纹传感器 814 采集到的指纹识别用户的身份, 或者, 由指纹传感器 814 根据采集到的指纹识别用户的身份。在识别出用户的身份为可信身份时, 由处理器 801 授权该用户执行相关的敏感操作, 该敏感操作包括解锁屏幕、查看加密信息、下载软件、

支付及更改设置等。指纹传感器 814 可以被设置终端 800 的正面、背面或侧面。当终端 800 上设置有物理按键或厂商 Logo 时，指纹传感器 814 可以与物理按键或厂商 Logo 集成在一起。

光学传感器 815 用于采集环境光强度。在一个实施例中，处理器 801 可以根据光学传感器 815 采集的环境光强度，控制触摸显示屏 805 的显示亮度。具体地，当环境光强度较高时，调高触摸显示屏 805 的显示亮度；当环境光强度较低时，调低触摸显示屏 805 的显示亮度。在另一个实施例中，处理器 801 还可以根据光学传感器 815 采集的环境光强度，动态调整摄像头组件 806 的拍摄参数。

接近传感器 816，也称距离传感器，通常设置在终端 800 的前面板。接近传感器 816 用于采集用户与终端 800 的正面之间的距离。在一个实施例中，当接近传感器 816 检测到用户与终端 800 的正面之间的距离逐渐变小时，由处理器 801 控制触摸显示屏 805 从亮屏状态切换为息屏状态；当接近传感器 816 检测到用户与终端 800 的正面之间的距离逐渐变大时，由处理器 801 控制触摸显示屏 805 从息屏状态切换为亮屏状态。

本领域技术人员可以理解，图 8 中示出的结构并不构成对终端 800 的限定，可以包括比图示更多或更少的组件，或者组合某些组件，或者采用不同的组件布置。

请参考图 9，其示出了本申请一个示例性实施例提供的服务器 900 的结构示意图。所述服务器 900 包括中央处理单元（CPU）901、包括随机存取存储器（RAM）902 和只读存储器（ROM）903 的系统存储器 904，以及连接系统存储器 904 和中央处理单元 901 的系统总线 905。所述服务器 900 还包括帮助计算机内的各个器件之间传输信息的基本输入/输出系统（I/O 系统）906，和用于存储操作系统 913、应用程序 914 和其他程序模块 915 的大容量存储设备 907。

所述基本输入/输出系统 906 包括有用于显示信息的显示器 908 和用于用户输入信息的诸如鼠标、键盘之类的输入设备 909。其中所述显示器 908 和输入设备 909 都通过连接到系统总线 905 的输入输出控制器 910 连接到中央处理单元 901。所述基本输入/输出系统 906 还可以包括输入输出控制器 910 以用于接收和处理来自键盘、鼠标、或电子触控笔等多个其他设备的输入。类似地，输入输出控制器 910 还提供输出到显示屏、打印机或其他类型的输出设备。

所述大容量存储设备 907 通过连接到系统总线 905 的大容量存储控制器(未示出)连接到中央处理单元 901。所述大容量存储设备 907 及其相关联的计算机可读介质为服务器 900 提供非易失性存储。也就是说,所述大容量存储设备 907 可以包括诸如硬盘或者 CD-ROM 驱动器之类的计算机可读介质(未示出)。

不失一般性,所述计算机可读介质可以包括计算机存储介质和通信介质。计算机存储介质包括以用于存储诸如计算机可读指令、数据结构、程序模块或其他数据等信息的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质。计算机存储介质包括 RAM、ROM、EPROM、EEPROM、闪存或其他固态存储其技术,CD-ROM、DVD 或其他光学存储、磁带盒、磁带、磁盘存储或其他磁性存储设备。当然,本领域技术人员可知所述计算机存储介质不局限于上述几种。上述的系统存储器 904 和大容量存储设备 907 可以统称为存储器。

根据本申请的各种实施例,所述服务器 900 还可以通过诸如因特网等网络连接到网络上的远程计算机运行。也即服务器 900 可以通过连接在所述系统总线 905 上的网络接口单元 911 连接到网络 912,或者说,也可以使用网络接口单元 911 来连接到其他类型的网络或远程计算机系统(未示出)。

可选的,该存储器中存储有至少一条指令、至少一段程序、代码集或指令集,至少一条指令、至少一段程序、代码集或指令集由处理器加载并执行以实现上述各个方法实施例提供的音视频处理方法。

本申请还提供一种计算机可读存储介质,该计算机可读存储介质存储有至少一条指令,至少一条指令用于被处理器执行以实现上述各个方法实施例提供的音视频处理方法。

本申请还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各个实施例所述的音视频处理方法。

上述本申请实施例序号仅仅为了描述,不代表实施例的优劣。

本领域普通技术人员可以理解实现上述实施例的音视频处理方法中全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。以上所述仅为本申请的较佳实施例,并不用以限制本申请,凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,

均应包含在本申请的保护范围之内。

权 利 要 求 书

1、一种音视频处理方法，其特征在于，用于终端中，所述方法包括：

显示目标音频的视频创建界面，所述视频创建界面中包括所述目标音频的 n 个音频片段和所述 n 个音频片段各自对应的视频录制入口，其中， $n \geq 2$ ；

在所述视频创建界面中接收作用于目标视频录制入口的触发信号，所述目标视频录制入口为与目标音频片段对应的视频录制入口；

根据所述触发信号获取所述目标音频片段对应的目标视频，所述目标视频为视频时长小于时长阈值的视频片段；

向所述服务器发送携带有所述目标视频的视频创建请求，所述视频创建请求用于指示在播放所述目标音频时播放所述目标视频的画面信息。

2、根据权利要求 1 所述的方法，其特征在于，所述根据所述触发信号获取所述目标音频片段对应的目标视频，包括：

根据所述触发信号开启视频录制功能，进行视频录制得到所述目标视频；

或，

根据所述触发信号获取所述终端中存储的所述目标视频；

或，

根据所述触发信号向所述服务器发送视频获取请求，所述视频获取请求中携带有所述目标音频片段的片段标识，所述视频获取请求用于指示所述服务器反馈所述片段标识对应的至少一个视频；从接收到的所述服务器反馈的所述至少一个视频中确定所述目标视频。

3、根据权利要求 1 所述的方法，其特征在于，所述显示目标音频的视频创建界面，包括：

显示所述目标音频的音频关联界面，所述音频关联界面中包括视频创建入口；

当接收到作用于视频创建入口的触发信号时，显示所述目标音频的视频创建界面。

4、根据权利要求 1 所述的方法，其特征在于，所述目标音频片段包括至少

两个所述目标音频片段，所述至少两个目标音频片段各自对应有所述目标视频；
所述向所述服务器发送携带有所述目标视频的视频创建请求之前，还包括：
将至少两个所述目标视频各自对应的画面信息进行拼接得到第一视频拼接画面；

所述向所述服务器发送携带有所述目标视频的视频创建请求，包括：

向所述服务器发送携带有所述第一视频拼接画面的第一视频创建请求，所述第一视频创建请求用于指示在播放所述目标音频时播放所述第一视频拼接画面。

5、根据权利要求1所述的方法，其特征在于，所述方法还包括：

向所述服务器发送所述目标音频的音频播放请求，所述音频播放请求用于指示所述服务器获取所述n个音频片段各自对应的第一候选视频并反馈画面播放指令，所述第一候选视频包括所述目标视频；

根据接收到的所述画面播放指令，在播放所述目标音频的音频片段时播放所述n个音频片段各自对应的所述第一候选视频的画面信息。

6、根据权利要求1至5任一所述的方法，其特征在于，所述方法还包括：

当接收到作用于第一音频片段的画面切换操作信号时，获取所述目标音频当前的第一播放时间点；

向所述服务器发送画面切换请求，所述画面切换请求携带有所述目标音频的音频标识，所述画面切换请求用于指示所述服务器获取与所述音频标识对应的第二视频拼接画面；向所述终端发送第一画面切换指令；

根据接收到的所述第一画面切换指令，在所述目标音频的第一播放时间点上开始切换播放所述第二视频拼接画面。

7、根据权利要求6所述的方法，其特征在于，所述画面切换请求中还携带有所述第一播放时间点，所述画面切换请求还用于指示所述服务器当不存在与所述音频标识对应的所述第二视频拼接画面时，获取与所述第一播放时间点所在的所述第一音频片段对应的第二候选视频；向所述终端发送第二画面切换指令；

所述方法还包括：

根据接收到的所述第二画面切换指令，在播放所述第一音频片段的过程中在所述目标音频的所述第一播放时间点上开始切换播放所述第二候选视频的画面信息。

8、根据权利要求5所述的方法，其特征在于，所述方法还包括：

在播放所述目标音频的指定时间内容时，将第一候选视频的画面信息切换显示为所述第一候选视频的视频制作信息，所述第一候选视频为所述指定时间内容所在的音频片段对应的候选视频；

所述指定时间内容包括前奏内容、间奏内容和尾奏内容中的至少一种，所述视频制作信息包括所述第一候选视频对应的录制者信息、上传者信息和拼接者信息中的至少一种。

9、一种音视频处理装置，其特征在于，用于终端中，所述装置包括：

显示模块，用于显示目标音频的视频创建界面，所述视频创建界面中包括所述目标音频的 n 个音频片段和所述 n 个音频片段各自对应的视频录制入口，其中， $n \geq 2$ ；

接收模块，用于在所述视频创建界面中接收作用于目标视频录制入口的触发信号，所述目标视频录制入口为与目标音频片段对应的视频录制入口；

获取模块，用于根据所述触发信号获取所述目标音频片段对应的目标视频，所述目标视频为视频时长小于时长阈值的视频片段；

发送模块，用于向所述服务器发送携带有所述目标视频的视频创建请求，所述视频创建请求用于指示在播放所述目标音频时播放所述目标视频的画面信息。

10、根据权利要求9所述的装置，其特征在于，所述获取模块，还用于根据所述触发信号开启视频录制功能，进行视频录制得到所述目标视频；

或，

根据所述触发信号获取所述终端中存储的所述目标视频；

或，

根据所述触发信号向所述服务器发送视频获取请求，所述视频获取请求中携带有所述目标音频片段的片段标识，所述视频获取请求用于指示所述服务器反馈所述片段标识对应的至少一个视频；从接收到的所述服务器反馈的所述至少一个视频中确定所述目标视频。

11、根据权利要求 9 所述的装置，其特征在于，所述显示模块，还用于显示所述目标音频的音频关联界面，所述音频关联界面中包括视频创建入口；

当接收到作用于视频创建入口的触发信号时，显示所述目标音频的视频创建界面。

12、根据权利要求 9 所述的装置，其特征在于，所述目标音频片段包括至少两个所述目标音频片段，所述至少两个目标音频片段各自对应有所述目标视频；

所述装置还包括：

拼接模块，用于将至少两个所述目标视频各自对应的画面信息进行拼接得到第一视频拼接画面；

发送模块，还用于向所述服务器发送携带有所述第一视频拼接画面的第一视频创建请求，所述第一视频创建请求用于指示在播放所述目标音频时播放所述第一视频拼接画面。

13、根据权利要求 9 所述的装置，其特征在于，所述装置还包括：

播放模块，用于向所述服务器发送所述目标音频的音频播放请求，所述音频播放请求用于指示所述服务器获取所述 n 个音频片段各自对应的第一候选视频并反馈画面播放指令，所述第一候选视频包括所述目标视频；根据接收到的所述画面播放指令，在播放所述目标音频的音频片段时播放所述 n 个音频片段各自对应的所述第一候选视频的画面信息。

14、根据权利要求 9 至 13 任一所述的装置，其特征在于，所述装置还包括：

第一画面切换模块，用于当接收到作用于第一音频片段的画面切换操作信号时，获取所述目标音频当前的第一播放时间点；

向所述服务器发送画面切换请求，所述画面切换请求携带有所述目标音频的音频标识，所述画面切换请求用于指示所述服务器获取与所述音频标识对应的第二视频拼接画面；向所述终端发送第一画面切换指令；

根据接收到的所述第一画面切换指令，在所述目标音频的第一播放时间点上开始切换播放所述第二视频拼接画面。

15、根据权利要求 14 所述的装置，其特征在于，所述画面切换请求中还携带有所述第一播放时间点，所述画面切换请求还用于指示所述服务器当不存在与所述音频标识对应的所述第二视频拼接画面时，获取与所述第一播放时间点所在的所述第一音频片段对应的第二候选视频；向所述终端发送第二画面切换指令；

所述装置还包括：

根据接收到的所述第二画面切换指令，在播放所述第一音频片段的过程中在所述目标音频的所述第一播放时间点上开始切换播放所述第二候选视频的画面信息。

16、根据权利要求 13 所述的装置，其特征在于，该装置还包括：

切换显示模块，用于在播放所述目标音频的指定时间内容时，将第一候选视频的画面信息切换显示为所述第一候选视频的视频制作信息，所述第一候选视频为所述指定时间内容所在的音频片段对应的候选视频；所述指定时间内容包括前奏内容、间奏内容和尾奏内容中的至少一种，所述视频制作信息包括所述第一候选视频对应的录制者信息、上传者信息和拼接者信息中的至少一种。

17、一种终端，其特征在于，所述终端包括处理器和存储器；所述存储器存储有至少一条指令，所述至少一条指令用于被所述处理器执行以实现如权利要求 1 至 8 任一所述的音视频处理方法。

18、一种计算机可读存储介质，其特征在于，所述存储介质存储有至少一条指令，所述至少一条指令用于被处理器执行以实现如权利要求 1 至 8 任一所述的音视频处理方法。