视频处理方法、装置、电子设备、存储介质

本申请要求于 2023 年 08 月 31 日提交中国专利局、申请号为 202311120475.X、发明名称为“视频处理方法、装置、电子设备及存储介质”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

本公开涉及视频技术领域，尤其涉及一种视频处理方法、装置、电子设备、存储介质及程序产品。

背景技术

随着互联网信息的日益丰富，观看视频已成为人们日常生活中的一项娱乐活动。为了提高用户的观看体验，在发布各类视频之前，对视频进行剪辑是一个重要环节。

创作者在创作时，可以使用视频剪辑工具，对拍摄好的视频素材进行剪辑。目前，大多数视频剪辑工具，主要是对视频帧的剪辑操作，包括对视频帧进行拖拽、排序、切换素材、加字幕等操作，进而实现对视频的剪辑过程。但是，对视频帧的剪辑操作一方面要求创作者具备一定的视频剪辑知识，另一方面创作者无法快速完成视频剪辑工作，因此，对视频帧的剪辑操作比较复杂且耗时，影响创作者的使用体验。

发明内容

为了解决上述技术问题或者至少部分地解决上述技术问题，本公开提供了一种视频处理方法、装置、电子设备、存储介质及程序产品。

本公开实施例的第一方面，提供一种视频处理方法，该方法包括：展示待处理视频对应的待处理文本数据；其中，待处理视频包括媒体视频和媒体音频，待处理文本数据对应的文本时间码信息，与媒体视频对应的第一时间线信息存在第一映射关系；文本时间码信息与媒体音频对应的第二时间线信息存在第二映射关系；响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段；基于第二映射关系，从媒体音频中确定目标文本数据对应的目标音频片段； 基于第一触发操作，针对目标文本数据、目标视频片段和目标音频片段进行相应处理。

本申请一些实施例中，待处理视频还包括至少一个包装对象，文

本时间码信息与每个包装对象对应的时间线信息存在映射关系；其中，每个包装对象的对象类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包、背景音乐。

本申请一些实施例中，第一触发操作包括以下任一项：删除操作、拖动操作、复制操作、剪切操作；该方法还包括：响应于针对待处理文本数据中的目标文本数据的第一触发操作，从至少一个包装对象中确定目标文本数据对应的目标包装对象；基于第一触发操作，在目标包装对象的对象类型为第一类型中的任一种的情况下，对目标包装对象进行相应处理；在目标包装对象的对象类型为第二类型中的任一种的情况下，禁止对目标包装对象进行相应处理；其中，第一类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包；第二类型包括：背景音乐。

本申请一些实施例中，第一类型包括第三类型和第四类型；在目标包装对象的对象类型为第一类型中的任一种的情况下，对目标包装对象进行相应处理，包括：在目标包装对象的对象类型为第三类型中的任一种的情况下，对目标包装对象中的目标包装片段进行相应处理，目标包装片段为目标包装对象中与目标文本数据对应的部分；或者，在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳在目标文本数据对应的时间码范围内的情况下，对目标包装对象的整体进行相应处理；其中，第三类型包括以下至少一项：字幕文本、画中画；第四类型包括以下至少一项：非字幕文本、贴纸、特效、滤镜、素材包。

本申请一些实施例中，该方法还包括：在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳不在目标文本数据对应的时间码范围内的情况下，禁止对目标包装对象进行相应处理。

本申请一些实施例中，该方法还包括：响应于针对待处理文本数据中的第一文本数据的修改操作，获取与修改操作对应的修改后的第二文本数据；基于第二映射关系，从媒体音频中确定与第一文本数据对应的第一音频片段；基于第二文本数据和第一音频片段的音色信息生成第二音频片段；将媒体音频中的第一音频片段替换为第二音频片段。

本申请一些实施例中，待处理视频还包括字幕文本，文本时间码信息与字幕文本对应的第三时间线信息存在第三映射关系；该方法还包括：响应于针对待处理文本数据中的第一文本数据的修改操作，基于第三映射关系，从字幕文本中确定与第一文本数据对应的第一字幕

片段；基于第二文本数据，对第一字幕片段进行修改，得到第二字幕片段；将字幕文本中的第一字幕片段替换为第二字幕片段。

本申请一些实施例中，第一触发操作还包括选中操作；响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段之前，该方法还包括：展示第一时间轴轨道和第二时间轴轨道，第一时间轴轨道为媒体视频对应的时间轴轨道，第二时间轴轨道为媒体音频对应的时间轴轨道；基于第一触发操作，针对目标文本数据、目标视频片段和目标音频片段进行相应处理，包括：按照第一显示样式，对目标文本数据进行显示；按照第二显示样式，对第一时间轴轨道中与目标视频片段对应的部分进行显示；按照第三显示样式，对第二时间轴轨道中与目标音频片段对应的部分进行显示。

本申请一些实施例中，该方法还包括：响应于针对待处理文本数据中的被选中文本数据的第二触发操作，在被选中文本数据对应的区域展示工具菜单，工具菜单包括以下至少一类：剪辑控件、添加控件，剪辑控件包括以下至少一项：与删除操作对应的删除控件、与复制操作对应的复制控件、与剪切控件对应的剪切控件、与修改操作对应的修改控件，智能长短句控件、搜索控件、替换控件；添加控件包括以下至少一项：添加画中画控件、添加非字幕文本控件、添加贴纸控件、添加特效控件、添加音效控件、添加滤镜控件、添加素材包控件、添加背景音乐控件；响应于对工具菜单中的目标控件的第三触发操作，对媒体视频、待处理文本数据、媒体音频中的至少一项执行相应处理。

本申请一些实施例中，在展示待处理视频对应的待处理文本数据之后，该方法还包括：响应于对素材库控件的触发操作，展示多个预设视频素材；响应于对多个预设视频素材中的目标视频素材的第五触发操作，展示文本粗剪浮窗，文本粗剪浮窗中包括目标视频素材对应的素材文本数据，目标视频素材包括目标素材视频和目标素材音频，素材文本数据的素材时间码信息与目标素材视频的时间线信息存在第四映射关系，素材时间码信息与目标素材音频的时间线信息存在第五映射关系；响应于对素材文本数据中的被选中素材文本的文本粗剪操作，将被选中素材文本与待处理文本数据合成，得到待剪辑文本数据，将素材视频片段与媒体视频合成为待剪辑媒体视频，将素材音频片段与媒体音频合成为待剪辑媒体音频，以使基于对待剪辑文本数据的文本剪辑操作，实现对待剪辑视频的剪辑；其中，素材视频片段为基于第四映射关系，从目标素材视频中获取的与被选中素材文本对应的视频片段；素材音频片段为基于第五映射关系，从目标素材音频中获取

的与被选中素材文本对应的音频片段；待剪辑视频包括待剪辑媒体视频和待剪辑媒体音频。

本公开实施例的第二方面，提供一种视频处理装置，该装置包括：展示模块，用于展示待处理视频对应的待处理文本数据；其中，待处理视频包括媒体视频和媒体音频，待处理文本数据对应的文本时间码信息，与媒体视频对应的第一时间线信息存在第一映射关系；文本时间码信息与媒体音频对应的第二时间线信息存在第二映射关系；确定模块，用于响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段；确定模块，还用于基于第二映射关系，从媒体音频中确定目标文本数据对应的目标音频片段； 处理模块，用于基于第一触发操作，针对目标文本数据、目标视频片段和目标音频片段进行相应处理。

本申请一些实施例中，待处理视频还包括至少一个包装对象，文本时间码信息与每个包装对象对应的时间线信息存在映射关系；其中，每个包装对象的对象类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包、背景音乐。

本申请一些实施例中，第一触发操作包括以下任一项：删除操作、拖动操作、复制操作、剪切操作；该确定模块，还用于响应于针对待处理文本数据中的目标文本数据的第一触发操作，从至少一个包装对象中确定目标文本数据对应的目标包装对象；该处理模块，还用于基于第一触发操作，在目标包装对象的对象类型为第一类型中的任一种的情况下，对目标包装对象进行相应处理；在目标包装对象的对象类型为第二类型中的任一种的情况下，禁止对目标包装对象进行相应处理；其中，第一类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包；第二类型包括：背景音乐。

本申请一些实施例中，第一类型包括第三类型和第四类型；该处理模块，具体用于在目标包装对象的对象类型为第三类型中的任一种的情况下，对目标包装对象中的目标包装片段进行相应处理，目标包装片段为目标包装对象中与目标文本数据对应的部分；或者，在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳在目标文本数据对应的时间码范围内的情况下，对目标包装对象的整体进行相应处理；其中，第三类型包括以下至少一项：字幕文本、画中画；第四类型包括以下至少一项：非字幕文本、贴纸、特效、滤镜、素材包。

本申请一些实施例中，该处理模块，还用于在目标包装对象的对

象类型为第四类型中的任一种，且目标包装对象的起始时间戳不在目标文本数据对应的时间码范围内的情况下，禁止对目标包装对象进行相应处理。

本申请一些实施例中，该装置还包括：获取模块、生成模块和替换模块，该获取模块，用于响应于针对待处理文本数据中的第一文本数据的修改操作，获取与修改操作对应的修改后的第二文本数据；该确定模块，还用于基于第二映射关系，从媒体音频中确定与第一文本数据对应的第一音频片段；该生成模块，用于基于第二文本数据和第一音频片段的音色信息生成第二音频片段；该替换模块，用于将媒体音频中的第一音频片段替换为第二音频片段。

本申请一些实施例中，待处理视频还包括字幕文本，文本时间码信息与字幕文本对应的第三时间线信息存在第三映射关系；该装置还包括：修改模块和替换模块；该确定模块，还用于响应于针对待处理文本数据中的第一文本数据的修改操作，基于第三映射关系，从字幕文本中确定与第一文本数据对应的第一字幕片段；该修改模块，用于基于第二文本数据，对第一字幕片段进行修改，得到第二字幕片段；该替换模块，用于将字幕文本中的第一字幕片段替换为第二字幕片段。

本申请一些实施例中，第一触发操作还包括选中操作；该展示模块，还用于在响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段之前，展示第一时间轴轨道和第二时间轴轨道，第一时间轴轨道为媒体视频对应的时间轴轨道，第二时间轴轨道为媒体音频对应的时间轴轨道；该处理模块，具体用于按照第一显示样式，对目标文本数据进行显示；按照第二显示样式，对第一时间轴轨道中与目标视频片段对应的部分进行显示；按照第三显示样式，对第二时间轴轨道中与目标音频片段对应的部分进行显示。

本申请一些实施例中，该展示模块，还用于在响应于针对待处理文本数据中的被选中文本数据的第二触发操作，在被选中文本数据对应的区域展示工具菜单，工具菜单包括以下至少一类：剪辑控件、添加控件，剪辑控件包括以下至少一项：与删除操作对应的删除控件、与复制操作对应的复制控件、与剪切控件对应的剪切控件、与修改操作对应的修改控件，智能长短句控件、搜索控件、替换控件；添加控件包括以下至少一项：添加画中画控件、添加非字幕文本控件、添加贴纸控件、添加特效控件、添加音效控件、添加滤镜控件、添加素材包控件、添加背景音乐控件；该处理模块，还用于响应于对工具菜单中的目标控件的第三触发操作，对媒体视频、待处理文本数据、媒体音频

中的至少一项执行相应处理。

本申请一些实施例中，该展示模块，用于在展示待处理视频对应的待处理文本数据之后，响应于对素材库控件的触发操作，展示多个预设视频素材；响应于对多个预设视频素材中的目标视频素材的第五触发操作，展示文本粗剪浮窗，文本粗剪浮窗中包括目标视频素材对应的素材文本数据，目标视频素材包括目标素材视频和目标素材音频，素材文本数据的素材时间码信息与目标素材视频的时间线信息存在第四映射关系，素材时间码信息与目标素材音频的时间线信息存在第五映射关系；响应于对素材文本数据中的被选中素材文本的文本粗剪操作，将被选中素材文本与待处理文本数据合成，得到待剪辑文本数据，将素材视频片段与媒体视频合成为待剪辑媒体视频，将素材音频片段与媒体音频合成为待剪辑媒体音频，以使基于对待剪辑文本数据的文本剪辑操作，实现对待剪辑视频的剪辑；其中，素材视频片段为基于第四映射关系，从目标素材视频中获取的与被选中素材文本对应的视频片段；素材音频片段为基于第五映射关系，从目标素材音频中获取的与被选中素材文本对应的音频片段；待剪辑视频包括待剪辑媒体视频和待剪辑媒体音频。

本公开实施例的第三方面，提供一种电子设备，该电子设备包括处理器、存储器及存储在该存储器上并可在该处理器上运行的计算机程序，该计算机程序被该处理器执行时实现如第一方面所述的视频处理方法。

本公开实施例的第四方面，提供一种计算机可读存储介质，该计算机可读存储介质上存储计算机程序，该计算机程序被处理器执行时实现如第一方面所述的视频处理方法。

本公开实施例的第五方面，提供了一种计算机程序产品，其中，该计算机程序产品包括计算机程序，当该计算机程序产品在处理器上运行时，使得处理器执行该计算机程序，实现如第一方面所述的视频处理方法。

本公开实施例的第六方面，提供了一种芯片，该芯片包括处理器和通信接口，该通信接口和该处理器耦合，该处理器用于运行程序指令，实现如第一方面所述的视频处理方法。

附图说明

此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本公开的实施例，并与说明书一起用于解释本公开的原理。

为了更清楚地说明本公开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，对于本领域普通技术人员而言，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本公开实施例提供的视频处理方法的流程示意图之一；

图 2 为本公开实施例提供的视频处理方法的流程示意图之二；

图 3 为本公开实施例提供的视频处理方法的界面示意图之一；

图 4 为本公开实施例提供的视频处理方法的界面示意图之二；

图 5 为本公开实施例提供的视频处理方法的界面示意图之三；

图 6 为本公开实施例提供的一种视频处理装置的结构框图；

图 7 为本公开实施例提供的一种电子设备的结构框图。

具体实施方式

为了能够更清楚地理解本公开的上述目的、特征和优点，下面将对本公开的方案进行进一步描述。需要说明的是，在不冲突的情况下，本公开的实施例及实施例中的特征可以相互组合。

在下面的描述中阐述了很多具体细节以便于充分理解本公开，但本公开还可以采用其他不同于在此描述的方式来实施；显然，说明书中的实施例只是本公开的一部分实施例，而不是全部的实施例。

本公开的说明书和权利要求书中的术语“第一”、“第二”等是用于区别类似的对象，而不用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便本公开的实施例能够以除了在这里图示或描述的那些以外的顺序实施，且“第一”、“第二”等所区分的对象通常为一类，并不限定对象的个数，例如第一对象可以是一个，也可以是多个。此外，说明书以及权利要求中“和/或”表示所连接对象的至少其中之一，字符“/”，一般表示前后关联对象是一种“或”的关系。

本公开实施例中的电子设备可以为移动电子设备，也可以为非移动电子设备。移动电子设备可以为手机、平板电脑、笔记本电脑、掌上电脑、车载电子设备、可穿戴设备、超级移动个人计算机（ultra-mobile personal computer，UMPC）、上网本或者个人数字助理（personal digital assistant，PDA ）等；非移动电子设备可以为个人计算机（personal computer，PC）、电视机（television，TV）、柜员机或者自助机等；本公开实施例不作具体限定。

本公开实施例提供的视频处理方法的执行主体可以为上述的电子设备（包括移动电子设备和非移动电子设备），也可以为该电子设备中能够实现该视频处理方法的功能模块和/或功能实体，具体的可以根据实际使用需求确定，本公开实施例不作限定。

下面结合附图，通过具体的实施例及其应用场景对本公开实施例提供的视频处理方法进行详细地说明。

如图 1 所示，本公开实施例提供一种视频处理方法，该方法可以包括下述的步骤 101 至步骤 104。

101、展示待处理视频对应的待处理文本数据。

其中，待处理视频包括媒体视频和媒体音频，待处理文本数据对应的文本时间码信息，与媒体视频对应的第一时间线信息存在第一映射关系；文本时间码信息与媒体音频对应的第二时间线信息存在第二映射关系。

其中，文本时间码信息包括待处理文本数据中每个语义字符对应的时间戳，第一时间线信息包括媒体视频的每个视频帧的时间戳。第一映射关系是待处理文本数据中每个语义字符的时间戳，与媒体视频帧的每个视频帧的时间戳的对应关系。示例性地，可以待处理文本数据中一个语义字符的时间戳对应一个视频帧的时间戳（即待处理文本数据中语义字符的时间戳与媒体视频中视频帧的时间戳一一对应），也可以待处理文本数据中一个语义字符的时间戳对应多个视频帧的时间戳。

其中，语义字符指具有单独语义的字符，例如，在中文中，一个语义字符指一个汉字，在英文中，一个语义字符至一个单词。

其中，第二时间线信息包括媒体音频中每个音频帧的时间戳。第二映射关系是待处理文本数据中每个语义字符的时间戳，与媒体音频的每个音频帧的时间戳的对应关系。示例性地，可以待处理文本数据中一个语义字符的时间戳对应一个音频帧的时间戳（即待处理文本数据中语义字符的时间戳与媒体音频中音频帧的时间戳一一对应），也可以待处理文本数据中一个语义字符的时间戳对应多个音频帧的时间戳。

可以理解，基于第一映射关系，可以从媒体视频中确定待处理文本数据中任意一个语义字符对应的至少一个视频帧，也可以从媒体视频中确定待处理文本数据中任意多个语义字符对应的多个视频帧。

可以理解，基于第二映射关系，可以从媒体音频中确定待处理文本数据中任意一个语义字符对应的至少一个音频帧，也可以从媒体音频中确定待处理文本数据中任意多个语义字符对应的多个音频帧。

本申请实施例中，待处理视频包括但不限于录制得到的视频、基于脚本得到的视频等。待处理文本数据可以是对媒体音频或待处理视频进行语音识别得到的，也可以是脚本。其中，待处理文本数据为脚本的情况下，可以将待处理文本数据与媒体视频匹配得到待处理文本数据与媒体视频的视频时间戳的第一映射关系，语音识别方法包括但不限于自动语音识别(Automatic Speech Recognition，ASR)技术。

本申请一些实施例中，待处理文本数据为脚本的情况下，可以将待处理文本数据与媒体音频匹配得到待处理文本数据与媒体音频的音频时间戳的第二映射关系；也可以根据媒体视频的视频时间戳与媒体音频的音频时间戳之间的映射关系，以及第一映射关系，获取第二映射关系；具体可以根据实际情况确定，此处不做限定。

本申请一些实施例中，可以通过响应于对展示待处理视频的界面中的文本剪辑控件的触发操作，调起文本剪辑面板，同时提取该待处理视频的媒体视频和媒体音频，然后对媒体音频进行语音识别得到待处理文本数据，在文本剪辑面板展示待处理文本数据，同时获取第一映射关系和第二映射关系。

本申请一些实施例中，在文本剪辑面板展示待处理文本数据的同时，还可以展示待处理视频、媒体视频对应的第一时间轴轨道、媒体音频对应的第二时间轴轨道、文本剪辑菜单、全局功能菜单中的至少一个。

其中，文本剪辑菜单可以包括以下至少一项：与删除操作对应的删除控件、与复制操作对应的复制控件、与剪切控件对应的剪切控件、与修改操作对应的修改控件，智能长短句控件（智能长短句替换控件）、替换控件、添加画中画控件、添加非字幕文本控件、添加贴纸控件、添加特效控件、添加音效控件、添加滤镜控件、添加素材包控件、添加背景音乐控件。

其中，全局功能菜单以下至少一项：搜索控件、去水词控件、字幕开关控件。

其中，去水词控件用于指示去水词处理，去水词处理是指去除待处理文本数据中的无意义词、语气词、静音片段等无效词汇或片段，以保证去水词处理后的待处理文本数据的语义的连贯性。

本申请一些实施例中，文本剪辑面板上还包括退出控件，响应于对退出控件的触发操作收起（退出）文本剪辑面板。

本申请一些实施例中，若在退出文本剪辑面时仍在对媒体音频进行 ASR 识别，识别不中断并在后台持续进行。

本实施例中，可在界面上展示待处理文本数据，然后通过对待处理文本数据的剪辑处理，实现对待处理视频的剪辑处理。

102、响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段。

第一触发操作可以为点击操作、滑动操作、拖动操作等，具体可以根据实际情况确定，此处不做限定。

可以理解，目标视频片段的时间戳与目标文本数据的时间戳对应。

103、基于第二映射关系，从媒体音频中确定目标文本数据对应的目标音频片段。

可以理解，目标音频片段的时间戳与目标文本数据的时间戳对应。

104、基于第一触发操作，针对目标文本数据、目标视频片段和目标音频片段进行相应处理。

可以理解，基于第一触发操作，针对目标文本数据、目标视频片段和目标音频片段进行与第一触发操作相应的处理。

本申请实施例中，基于第一映射关系和第二映射关系，通过第一触发操作对待处理文本数据进行相应处理，实现对媒体视频和媒体音频进行相应处理，进而实现对待处理视频的剪辑处理。相比于视频帧的剪辑操作，对文本数据的剪辑操作相对简单，且对创作者的专业性要求较低，可以提高对待处理视频的剪辑效率，提高创作者的使用体验。

本申请一些实施例中，待处理视频还包括至少一个包装对象，文本时间码信息与每个包装对象对应的时间线信息存在映射关系；其中，每个包装对象的对象类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包、背景音乐。

包装对象可以对应完整的待处理文本数据（如包装对象为背景音乐、字幕文本、画中画等），包装对象也可以对应待处理文本数据中的部分文本数据（如包装对象为贴纸、特效、滤镜等）；具体可以根据实际情况确定，此处不做限定。

以至少一个包装对象包括的包装对象 1 和包装对象 2 为例，文本时间码信息与包装对象 1 对应的时间线信息存在映射关系 1，文本时间码信息与包装对象 2 对应的时间线信息存在映射关系 2。

其中，字幕文本是与待处理视频的内容、媒体音频的内容对应的文本。非字幕文本为除字幕文本之外的文本。

其中，画中画是一种视频内容呈现方式，是指在一部视频全屏播出的同时，于画面的小面积区域上同时播出另一部视频，被广泛用于电视、视频录像、监控、演示设备等等。

其中，贴纸可以包括文字、图片、图像等。

其中，特效可以包括视觉特效、声音特效（简称音效）等。

其中，滤镜，主要是用来实现图像的各种特殊效果，具体可以参考现有相关技术。

其中，素材包就像一个个组合好的小模板，它包含了画面、字幕、特效等元素，支持一键添加，并且在添加后可以根据需要更改其中的

元素。

其中，背景音乐又称伴乐，配乐；通常是指在电视剧、电影、动画、电子游戏、网站中用于调节气氛的一种音乐，插入于对话之中，能够增强情感的表达，达到一种让观众身临其境的感受。本申请实施例中背景音乐是与媒体音频、声音特效不同的音频内容。

本申请实施例中，待处理视频包括至少一个包装对象，且每个包装对象的时间线信息与待处理文本数据的文本时间码信息存在映射关系，因此在对待处理文本数据中的部分文本数据进行剪辑操作的过程中，可能需要对与部分文本数据匹配的包装对象进行相应处理。本申请实施例将待处理视频中的媒体视频、媒体音乐、至少一个包装对象等分别提取出来，然后根据对待处理文本数据的剪辑操作，实现针对媒体视频、媒体音乐、至少一个包装对象中的至少一个进行相应操作，以实现对待处理视频的剪辑，相对于基于对待处理文本数据的剪辑，对待处理视频整体进行剪辑，使得对待处理视频的剪辑更细节化和合理化，而且无需用户操作简单便于理解，可以提高用户体验。

本申请一些实施例中，第一触发操作包括以下任一项：删除操作、拖动操作、复制操作、剪切操作；第一触发操作还可以包括其他的操作，此处不做限定。本申请实施例中，第一触发操作可以支持多种剪辑操作，用户操作简单便于理解，可以提高用户体验。

本公开实施例提供的技术方案与现有技术相比具有如下优点：展示待处理视频对应的待处理文本数据；其中，待处理视频包括媒体视频和媒体音频，待处理文本数据对应的文本时间码信息，与媒体视频对应的第一时间线信息存在第一映射关系；文本时间码信息与媒体音频对应的第二时间线信息存在第二映射关系；响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段；基于第二映射关系，从媒体音频中确定目标文本数据对应的目标音频片段； 基于第一触发操作，针对目标文本数据、目标视频片段和目标音频片段进行相应处理。本申请实施例中，基于第一映射关系和第二映射关系，通过第一触发操作对待处理文本数据进行相应处理，实现对媒体视频和媒体音频进行相应处理，进而实现对待处理视频的剪辑处理。相比于视频帧的剪辑操作，对文本数据的剪辑操作相对简单，且对创作者的专业性要求较低，可以提高对待处理视频的剪辑效率，提高创作者的使用体验。

本申请一些实施例中，结合图 1，如图 2 所示，在上述步骤 101 之后，本申请实施例提供的视频处理方法还可以包括下述步骤 105 至

步骤 107。

105、响应于针对待处理文本数据中的目标文本数据的第一触发操作，从至少一个包装对象中确定目标文本数据对应的目标包装对象。

可以理解，目标包装对象为至少一个包装对象中，时间戳与目标文本数据的时间戳存在重合的包装对象。

106、基于第一触发操作，在目标包装对象的对象类型为第一类型中的任一种的情况下，对目标包装对象进行相应处理。

其中，第一类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包。

107、在目标包装对象的对象类型为第二类型中的任一种的情况下，禁止对目标包装对象进行相应处理。

其中，第二类型包括：背景音乐。

可以理解，本申请实施例中，对包装对象的类型进行区分，有些类型的包装对象与目标文本数据的剪辑联动，有些类型的包装对象不响应目标文本数据的剪辑操作。如此，可以根据包装对象的类别，确定是否对包装对象进行联动剪辑。

可以理解，在包装对象的类型为背景音乐，如果与目标文本数据进行联动剪辑，则背景音乐可能出现不连贯、重复等现象，不能够再起到增强情感，让观众身临其境的感受的效果。

本申请一些实施例中，第一类型包括第三类型和第四类型；上述步骤 106 具体可以通过下述步骤 106a 或步骤 106b 实现。

106a、在目标包装对象的对象类型为第三类型中的任一种的情况下，对目标包装对象中的目标包装片段进行相应处理，目标包装片段为目标包装对象中与目标文本数据对应的部分。

其中，第三类型包括以下至少一项：字幕文本、画中画。

本申请实施例中，字幕文本和画中画类型（第三类型）的包装对象一般是与媒体视频一一对应的，因此，可以与目标文本数据进行联动剪辑。

106b、在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳在目标文本数据对应的时间码范围内的情况下，对目标包装对象的整体进行相应处理。

其中，时间码分为为目标文本数据中的第一个语义字符对应的时间码（时间戳）到最后一个语义字符对应的时间码之间的时间范围。

可以理解，目标包装对象的起始时间戳在目标文本数据对应的时间码范围内，即为目标包装对象的起始时间戳与时间码范围中的某一时间戳相同。

其中，第四类型包括以下至少一项：非字幕文本、贴纸、特效、滤镜、素材包。

本申请实施例中，非字幕文本、贴纸、特效、滤镜、素材包类型（第四类型）的包装对象通常情况下仅与媒体视频的一部分对应，如果不做确定是否与目标文本数据进行联动剪辑的判断，直接与目标文本数据进行联动，则可能导致包装对象不连贯，一个包装对象变多个等现象发生，丧失了包装对象原有的效果，反而增加了不好的效果。因此，本申请实施例中，在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳在目标文本数据对应的时间码范围内的情况下，对目标包装对象的整体进行相应处理，即在确定与目标文本数据进行联动剪辑的情况下，将包装对象整体进行相应处理，如此，可以保证在确定与目标文本数据进行联动剪辑的情况下，包装对象整体进行相应处理，不会增加不好的效果。

本申请一些实施例中，在上述步骤 105 之后，本申请实施例提供的视频处理方法还可以包括下述步骤 108。

108、在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳不在目标文本数据对应的时间码范围内的情况下，禁止对目标包装对象进行相应处理。

可以理解，在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳不在目标文本数据对应的时间码范围内的情况下，禁止对目标包装对象进行相应处理，即确定不响应与目标文本数据的联动剪辑（不与目标文本数据进行联动剪辑），如此，可以保证包装对象仍具有原有的效果。

本申请实施例中，修改操作作用于待处理文本数据和媒体音频，如有字幕文本也会联动，不作用于媒体视频。即通常情况下针对修改操作，媒体音频与待处理文本数据进行联动剪辑，如果待处理视频包括字幕文本，则字幕文本与待处理文本数据进行联动剪辑，而媒体视频不与待处理文本数据进行联动剪辑。如此，当对待处理文本数据进行修改时，使得媒体音频和字幕文本相应修改，但是无需修改媒体视频，逻辑简单，且贴合实际情况，用户体验较好。

本申请一些实施例中，在上述步骤 101 之后，本申请实施例提供的视频处理方法还可以包括下述步骤 109 至步骤 112。

109、响应于针对待处理文本数据中的第一文本数据的修改操作，获取与修改操作对应的修改后的第二文本数据。

110、基于第二映射关系，从媒体音频中确定与第一文本数据对应的第一音频片段。

111、基于第二文本数据和第一音频片段的音色信息生成第二音频片段。

本申请实施例中，基于第二文本数据和第一音频片段的音色信息生成第二音频片段，一方面可以保证第二音频片段的内容与第二文本数据对应，一方面又可以保证第二音频片段的音色信息与第一音频片段的音色信息相同，如此，不仅相对待处理文本数据的修改相应修改了媒体音频，而且保证了媒体音频中的修改位置的音色信息不发生变化，可以提高用户体验。

112、将媒体音频中的第一音频片段替换为第二音频片段。

可以理解，通过将媒体音频中的第一音频片段替换为第二音频片段，实现了媒体音频与第一文本数据的联动修改。

本申请一些实施例中，待处理视频还包括字幕文本，文本时间码信息与字幕文本对应的第三时间线信息存在第三映射关系；在上述步骤 109 之后，本申请实施例提供的视频处理方法还可以包括下述步骤 113 至步骤 115。

113、响应于针对待处理文本数据中的第一文本数据的修改操作，基于第三映射关系，从字幕文本中确定与第一文本数据对应的第一字幕片段。

114、基于第二文本数据，对第一字幕片段进行修改，得到第二字幕片段。

可以理解，第二字幕片段与第二文本数据对应，第二音频片段也与第二文本数据对应，因此，第二字幕片段也与第二音频片段对应，如此在观众观看时，不会发生字幕与音频不对应的情况。

115、将字幕文本中的第一字幕片段替换为第二字幕片段。

可以理解，通过将字幕文本中的第一字幕片段替换为第二字幕片段，实现了字幕文本与第一文本数据的联动修改。

本申请一些实施例中，在不显示媒体视频时间轴轨道、媒体音频时间轴轨道的情况下，选中操作只针对文本数据，被选中的文本数据进行选中显示。

本申请一些实施例中，在显示媒体视频时间轴轨道、媒体音频时间轴轨道的情况下，选中操作针对被选中的文本数据进行选中显示、针对媒体视频时间轴轨道中与被选中的文本数据对应的部分轨道进行选中显示、媒体音频时间轴轨道中与被选中的文本数据对应的部分轨道进行选中显示。如此，可以给用户更好的视觉体验，让用户明白对应的剪辑过程是如何实现的。

本申请一些实施例中，第一触发操作还包括选中操作；在上述步

骤 102 之前，本申请实施例提供的视频处理方法还可以包括下述步骤 116，上述步骤 104 具体可以通过下述步骤 104a 至 104c 实现。

116、展示第一时间轴轨道和第二时间轴轨道。

其中，第一时间轴轨道为媒体视频对应的时间轴轨道，第二时间轴轨道为媒体音频对应的时间轴轨道。

104a、按照第一显示样式，对目标文本数据进行显示。

104b、按照第二显示样式，对第一时间轴轨道中与目标视频片段对应的部分进行显示。

104c、按照第三显示样式，对第二时间轴轨道中与目标音频片段对应的部分进行显示。

其中，第一显示样式、第二显示样式和第三显示样式中可以部分相同，或部分不同，也可以均相同，或均不相同，此处不做限定。

本申请一些实施例中，在除了显示媒体视频时间轴轨道、媒体音频时间轴轨道之外，还显示每个包装对象的时间轴轨道的情况下，根据包装对象的类型，以及包装对象的起始时间戳是否在被选中文本数据对应的时间码范围，确定包装对象的时间轴轨道是否进行选中显示。

本申请实施例中的剪辑时轨道联动逻辑（具体与上述的媒体视频、媒体音频、包装对象是否与目标文本数据进行联动剪辑对应）包括：

先判断被选中文本的起始时间码和结束时间码，确定被选中文本的目标时间码范围；

媒体视频的时间轴轨道、媒体音频的时间轴轨道、画中画的时间轴轨道、文本字幕的时间轴轨道中分别与目标时间码范围对应的轨道部分进行相应的选中显示，且响应后续的删除、移动、复制等剪辑操作；

非文本字幕的时间轴轨道、贴纸对应的时间轴轨道等中的起始时间戳如果在目标时间码范围内，则对应的轨道整体进行相应的选中显示，且响应后续的删除、移动、复制等剪辑操作；

背景音乐的时间轴轨道始终不进行相应的选中显示，也不响应后续的删除、移动、复制等剪辑操作。

本申请一些实施例中，上述剪辑时轨道联动逻辑是默认的，也可以对不同的时间轴轨道设置联动开关控件，用户可以通过对联动开关控件的触发操作，确定是否对相应的时间轴轨道进行联动，此处不做限定。

本申请一些实施例中，本申请实施例提供的方法还可以包括下述的步骤 201 至步骤 202。

201、响应于针对待处理文本数据中的被选中文本数据的第二触发

操作，在被选中文本数据对应的区域展示工具菜单。

其中，工具菜单包括以下至少一类：剪辑控件、添加控件，剪辑控件包括以下至少一项：与删除操作对应的删除控件、与复制操作对应的复制控件、与剪切控件对应的剪切控件、与修改操作对应的修改控件，智能长短句控件、搜索控件、替换控件；添加控件包括以下至少一项：添加画中画控件、添加非字幕文本控件、添加贴纸控件、添加特效控件、添加音效控件、添加滤镜控件、添加素材包控件、添加背景音乐控件。

其中，上述任意一个添加控件可以支持用户手动选择添加或智能匹配添加。

202、响应于对工具菜单中的目标控件的第三触发操作，对媒体视频、待处理文本数据、媒体音频中的至少一项执行相应处理。

本申请实施例中，通过对被选中文本数据的第二触发操作，可以展示工具菜单，然后可以根据实际需求，对工具菜单中的任意控件的触发操作，执行相应的剪辑处理，交互操作简单便捷，有利于提高用户体验。

本申请一些实施例中，待处理文本数据所在的显示区域包括素材库控件，可对全局视频进行文本剪辑，也可对素材库内容进行粗剪。

本申请一些实施例中，在上述步骤 101 之后，本申请实施例提供的方法还可以包括下述的步骤 301 至步骤 305。

301、响应于对素材库控件的触发操作，展示多个预设视频素材。

302、响应于对多个预设视频素材中的目标视频素材的第五触发操作，展示文本粗剪浮窗。

其中，文本粗剪浮窗中包括目标视频素材对应的素材文本数据，目标视频素材包括目标素材视频和目标素材音频，素材文本数据的素材时间码信息与目标素材视频的时间线信息存在第四映射关系，素材时间码信息与目标素材音频的时间线信息存在第五映射关系。

303、响应于对素材文本数据中的被选中素材文本的文本粗剪操作，将被选中素材文本与待处理文本数据合成，得到待剪辑文本数据。

304、将素材视频片段与媒体视频合成为待剪辑媒体视频。

305、将素材音频片段与媒体音频合成为待剪辑媒体音频，以使基于对待剪辑文本数据的文本剪辑操作，实现对待剪辑视频的剪辑。

其中，素材视频片段为基于第四映射关系，从目标素材视频中获取的与被选中素材文本对应的视频片段；素材音频片段为基于第五映射关系，从目标素材音频中获取的与被选中素材文本对应的音频片段；待剪辑视频包括待剪辑媒体视频和待剪辑媒体音频。

本申请实施例中，文本剪辑面板中还可以包括素材库控件，通过对素材库控件的触发操作，可以展示多个预设视频素材，用户可以根据需求从多个预设视频素材中选择需要的目标素材视频，然后通过对目标素材视频的文本粗剪操作，拉起文本粗剪浮窗，粗剪浮窗中展示有目标素材视频对应的素材文本数据，因此用户可以通过对素材文本数据中的被选中素材文本的文本粗剪操作，剪辑目标素材视频中用户需要的剪辑素材片段，得到素材视频片段和素材音频片段，然后将被选中素材文本与待处理文本数据合成，得到待剪辑文本数据，将素材视频片段与媒体视频合成为待剪辑媒体视频，将素材音频片段与媒体音频合成为待剪辑媒体音频，以使基于对待剪辑文本数据的文本剪辑操作，实现对待剪辑视频的剪辑。

本申请实施例中，待处理视频也可以通过从素材库中选中至少一个素材，然后通过文本粗剪，得到的多个素材剪辑片段合成的，具体可以根据实际情况确定，此处不做限定。

本申请一些实施例中，可以选中素材库任意素材，右键菜单，新增文本粗剪入口。点击文本粗剪入口，呼出文本粗剪浮窗，同时开始对选中的素材进行 ASR 识别，支持在文本粗剪浮窗粗剪素材以添加到时间线。

文本粗剪浮窗包括浮窗文本预览区，还可以包括浮窗视频预览区、浮窗时间轴展示区等，具体可以参考文本剪辑面板的相关描述，并且可以根据实际情况确定，此处不做限定。

针对选中素材的 ASR 识别结果（素材文本数据）会展示在浮窗文本预览区。浮窗文本预览区展示素材文本数据的具体形式可以参考上述文本剪辑面板的文本预览区展示待处理文本数据的具体形式，此处不予赘述。

文本粗剪操作：选中素材文本，拖动选中的素材文本，进行粗剪操作；或者，选中素材文本，邮件调出粗剪控件，点击粗剪控件；具体可以根据实际情况确定，此处不做限定。

本申请一些实施例中，本地素材智能搜索关键词，选中搜索结果并进入文本粗剪时，默认匹配到对应文本并选中收起可快速完成粗剪。

针对输入文本数据支持通过剪辑浮窗进行文本粗剪的功能。可以通过对输入文本数据或匹配视频的触发操作，调出文本粗剪控件，然后通过对文本粗剪控件的触发操作，呼出新模态窗口，支持在此面板粗剪输入文本数据，并实现对匹配视频的相应剪辑，具体过程可以参考上述对待处理文本数据的剪辑操作，实现对待处理视频的剪辑，此处不予赘述。

本申请实施例中，通过剪辑浮窗实现对输入文本数据对应的匹配视频的剪辑，将剪辑处理后的待处理视频与剪辑处理后的匹配视频进行合成处理，得到目标剪辑视频。

本申请一些实施例中，不同展示对象可以在文本剪辑面板中分区域展示。例如，文本剪辑面板可以由以下中的部分或全部区域构成：视频预览区、文本预览区、文本剪辑菜单区、全局功能区、时间轴轨道展示区。其中，视频预览区用于展示待处理视频，文本预览区，用于展示待处理文本数据，文本剪辑菜单区用于展示文本剪辑菜单，全局功能区用于展示全局功能菜单。

本申请一些实施例中，在拉起文本剪辑面板时，采用默认布局模式。然后支持通过选择预设布局模式（如“工作模式”和“面板视图模式”等模式菜单）确定文本剪辑面板中可以展示的分区，也支持独立选择各个分区是否展示（勾选->显示；取消勾选->隐藏。例如，每个面板分区可通过视图菜单中的「√/x」展示/取消展示。）。而且面板分区可被拖动、使其悬浮在编辑器外。拖动还原时面板分区只能被回归到默认布局下的对应位置（即无法随意吸附）。

文本剪辑面板中还包括“还原”、“全屏”、“最小化”、“关闭”选项。

其中，文本预览区包括待处理文本数据和待处理文本数据对应的文本时间码信息，以及时间轴光标。待处理文本数据包括至少一个用户的语音对应的文本数据，待处理文本数据包括至少一个文本段落（包含含无声片段【 ---】）。时间轴光标：用于指示待处理视频对应的当前播放进度，对应待处理视频的当前播放时间轴，对应文本词或文本句会高亮。待处理视频播放时实时更新时间轴光标位置和高亮文本词或文本句。

本申请实施例中的文本分段逻辑可以根据以下优先级顺序进行分段：文本手动回行分段的优先级大于发言人分段的优先级，发言人分段的优先级大于按照停顿时长分段的优先级，按照停顿时长分段的优先级大于按照段落字数分段的优先级。

本申请实施例中的无声片段逻辑可以包括：方案 1：大于或等于 1s 的静音片段视为无声片段，在文本剪辑面板的文本预览区中展示；方案 2：大于或等于语速 10 倍的静音片段视为无声片段，在文本剪辑面板的文本预览区中展示。

本申请实施例中的定位跳转逻辑可以为手动滚动文本区域无响应，仅在点击切换选中时跳转。

本申请实施例中，可以通过对选中文本的触发操作调出工具菜单，

然后对工具菜单中的任意剪辑控件的触发操作或者直接通过对文本剪辑菜单中的任意剪辑控件的触发操作，进入文本剪辑模式，进行文本兼视频的剪辑。

待处理文本数据处于预览态（是默认状态）。此时播放时间点有光标且当前朗读文本词或文本句（不论暂停或播放中）高亮显示。此时，可以通过触发操作进入编辑态以修改内容。例如，可以通过对选择控件的触发操作，当前高亮句被选中，并且进入编辑态。

待处理文本数据中存在选中文本时进入编辑态，可进行如下操作：删除操作、插入 xxx（如文字模板/画中画等）操作、修改操作、移动操作、复制操作等。

删除操作，选中文本时，呼出工具菜单，可见删除控件，选择删除控件，删除选中文本和与与所述选中文本对应的轨道内容（具体删除的轨道取决于轨道联动逻辑），删除结束回到预览态。

插入 xxx（如文字模板/画中画等）操作，选中文本时，呼出工具菜单，可见添加 xxx 控件，拉起对应面板展示可以添加的包装对象的可选素材，以供用户自选，也可以根据选中文本智能匹配需要添加的包装对象的素材，具体可以根据实际情况确定，此处不做限定。

用户自选方式选择任意素材后或者智能匹配素材后，素材将以新轨道覆盖在选中文本的起始时间码。添加后，添加结果不在文本预览区以任意形式展示，但视频预览区（视频预览窗口，视频播放器）可见，故只可上屏编辑，插入结束仍处于编辑态。

移动操作，选中文本后，长按选中区域，可拖动选中的文本段落至任意位置（不可至词语内）、对应视频片段也会跟随移动，移动结束仍处于编辑态。

复制操作，选中文本时，呼出工具菜单，可见复制控件，选择复制控件，复制文本和轨道内容（具体复制的轨道取决于轨道联动开关），并粘贴到选中文本后，复制结束仍处于编辑态。

修改操作，此操作作用于待处理文本数据和媒体音频，如有字幕文本也会联动，不作用于媒体视频。选中文本时，呼出工具菜单，可见修改控件，选择后拉起输入框，待处理文本数据全文进入输入框，且当前被选中句也默认被选中。此修改操作的结果只作用于待处理文本数据、媒体音频（字幕文本）不作用于媒体视频，故修改操作响应结果如下：换行-字幕分两句；去掉回行-字幕合并为一句；修改错别字，字幕对应修改，有时对应修改媒体音频发音，输入新文本-视频时间码不会延长，文本修改结果只作用于字幕。选择输入框上的完成控件，结束修改操作，回到预览态，光标位于最后输入位置。

本申请实施例中，文本剪辑模式下，可进行以下全局功能操作：搜索操作、去水词操作、字幕开关操作。

搜索操作，进入文本剪辑模式后、选择搜索控件，展开搜索框，输入搜索关键词，点击确认，匹配到的文本以高亮标记，支持选择上一个控件和下一个控件，以快速跳转搜索结果。点击搜索框中的关闭控件，收起搜索框。

去水词操作，进入文本剪辑模式后、选择去水词控件，开始识别水词，去水词识别完成后、文本预览区中水词以高亮标记，并在底部新增确认（去水词）控件，可通过对确认控件的触发操作实现一键去水词功能，用户可确认水词识别结果是否准确，如不准确可选中单个水词呼出恢复控件（取消对应的高亮标记），如确认可选择确认控件以继续后续操作。选择确认控件，删除高亮标记出的所有水词并回到文本剪辑模式。

本申请一些实施例中，去水词操作会触发全文去水。如已处理过去水，二次点击会响应为再处理一遍。

字幕开关操作，字幕内容与文本内容联动，字幕开关控件（可以包括开启字幕控件和关闭字幕控件）只控制字幕是否可见。选择开启字幕控件，添加字幕，且文本数据更新时字幕必然联动更新；选择关闭字幕控件，不添加字幕。

示例 1，由于移动电子设备的展示尺寸有限，以及移动电子设备大多对应非专业用户，因此对于移动电子设备，文本剪辑面板可以默认展示如图 3 所示的文本预览区（对应展示待处理文本数据）、视频预览区（对应展示待处理视频）、文本剪辑菜单区（对应展示文本剪辑菜单）和全局功能区（对应展示全局功能菜单），其中，标记“31”指示的区域为视频预览区，标记“32”指示的区域为文本预览区，标记“33”指示的区域为全局功能区，标记“34”指示的区域为文本剪辑菜单区。其中，第一时间轴轨道、第二时间轴轨道、以及待处理视频中的其他包装对象对应的时间轴轨道，可以根据使用需求通过相应的触发操作，控制在文本剪辑面板进行相应展示，此处不做限定。

示例 2，在非移动电子设备可以默认展示如图 4 所示文本预览区、视频预览区、时间轴轨道展示区的三分布局模式，也可以默认展示如图 5 所示的文本预览区、视频预览区、时间轴轨道展示区，调参面板展示区的四分布局模式。其中，标记“41”指示的区域为视频预览区，标记“42”指示的区域为文本预览区，标记“43”指示的区域为时间轴轨道展示区，标记“51”指示的区域为调参面板展示区。

图 6 为本公开实施例示出的一种视频处理装置的结构框图，如图 6

所示，包括：

展示模块601，用于展示待处理视频对应的待处理文本数据；其中，待处理视频包括媒体视频和媒体音频，待处理文本数据对应的文本时间码信息，与媒体视频对应的第一时间线信息存在第一映射关系；文本时间码信息与媒体音频对应的第二时间线信息存在第二映射关系；

确定模块602，用于响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段；确定模块602，还用于基于第二映射关系，从媒体音频中确定目标文本数据对应的目标音频片段；

处理模块603，用于基于第一触发操作，针对目标文本数据、目标视频片段和目标音频片段进行相应处理。

本申请一些实施例中，待处理视频还包括至少一个包装对象，文本时间码信息与每个包装对象对应的时间线信息存在映射关系；其中，每个包装对象的对象类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包、背景音乐。

本申请一些实施例中，第一触发操作包括以下任一项：删除操作、拖动操作、复制操作、剪切操作；该确定模块602，还用于响应于针对待处理文本数据中的目标文本数据的第一触发操作，从至少一个包装对象中确定目标文本数据对应的目标包装对象；该处理模块603，还用于基于第一触发操作，在目标包装对象的对象类型为第一类型中的任一种的情况下，对目标包装对象进行相应处理；在目标包装对象的对象类型为第二类型中的任一种的情况下，禁止对目标包装对象进行相应处理；其中，第一类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包；第二类型包括：背景音乐。

本申请一些实施例中，第一类型包括第三类型和第四类型；该处理模块603，具体用于在目标包装对象的对象类型为第三类型中的任一种的情况下，对目标包装对象中的目标包装片段进行相应处理，目标包装片段为目标包装对象中与目标文本数据对应的部分；或者，在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳在目标文本数据对应的时间码范围内的情况下，对目标包装对象的整体进行相应处理；其中，第三类型包括以下至少一项：字幕文本、画中画；第四类型包括以下至少一项：非字幕文本、贴纸、特效、滤镜、素材包。

本申请一些实施例中，该处理模块603，还用于在目标包装对象的对象类型为第四类型中的任一种，且目标包装对象的起始时间戳不在目标文本数据对应的时间码范围内的情况下，禁止对目标包装对象进

行相应处理。

本申请一些实施例中，该装置还包括：获取模块、生成模块和替换模块，该获取模块，用于响应于针对待处理文本数据中的第一文本数据的修改操作，获取与修改操作对应的修改后的第二文本数据；该确定模块602，还用于基于第二映射关系，从媒体音频中确定与第一文本数据对应的第一音频片段；该生成模块，用于基于第二文本数据和第一音频片段的音色信息生成第二音频片段；该替换模块，用于将媒体音频中的第一音频片段替换为第二音频片段。

本申请一些实施例中，待处理视频还包括字幕文本，文本时间码信息与字幕文本对应的第三时间线信息存在第三映射关系；该装置还包括：修改模块和替换模块；该确定模块602，还用于响应于针对待处理文本数据中的第一文本数据的修改操作，基于第三映射关系，从字幕文本中确定与第一文本数据对应的第一字幕片段；该修改模块，用于基于第二文本数据，对第一字幕片段进行修改，得到第二字幕片段；该替换模块，用于将字幕文本中的第一字幕片段替换为第二字幕片段。

本申请一些实施例中，第一触发操作还包括选中操作；该展示模块601，还用于在响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段之前，展示第一时间轴轨道和第二时间轴轨道，第一时间轴轨道为媒体视频对应的时间轴轨道，第二时间轴轨道为媒体音频对应的时间轴轨道；该处理模块603，具体用于按照第一显示样式，对目标文本数据进行显示；按照第二显示样式，对第一时间轴轨道中与目标视频片段对应的部分进行显示；按照第三显示样式，对第二时间轴轨道中与目标音频片段对应的部分进行显示。

本申请一些实施例中，该展示模块601，还用于在响应于针对待处理文本数据中的被选中文本数据的第二触发操作，在被选中文本数据对应的区域展示工具菜单，工具菜单包括以下至少一类：剪辑控件、添加控件，剪辑控件包括以下至少一项：与删除操作对应的删除控件、与复制操作对应的复制控件、与剪切控件对应的剪切控件、与修改操作对应的修改控件，智能长短句控件、搜索控件、替换控件；添加控件包括以下至少一项：添加画中画控件、添加非字幕文本控件、添加贴纸控件、添加特效控件、添加音效控件、添加滤镜控件、添加素材包控件、添加背景音乐控件；该处理模块603，还用于响应于对工具菜单中的目标控件的第三触发操作，对媒体视频、待处理文本数据、媒体音频中的至少一项执行相应处理。

本申请一些实施例中，该装置还包括：合成模块；该展示模块601，

用于在展示待处理视频对应的待处理文本数据之后，响应于对素材库控件的触发操作，展示多个预设视频素材；响应于对多个预设视频素材中的目标视频素材的第五触发操作，展示文本粗剪浮窗，文本粗剪浮窗中包括目标视频素材对应的素材文本数据，目标视频素材包括目标素材视频和目标素材音频，素材文本数据的素材时间码信息与目标素材视频的时间线信息存在第四映射关系，素材时间码信息与目标素材音频的时间线信息存在第五映射关系；该合成模块，用于响应于对素材文本数据中的被选中素材文本的文本粗剪操作，将被选中素材文本与待处理文本数据合成，得到待剪辑文本数据，将素材视频片段与媒体视频合成为待剪辑媒体视频，将素材音频片段与媒体音频合成为待剪辑媒体音频，以使基于对待剪辑文本数据的文本剪辑操作，实现对待剪辑视频的剪辑；其中，素材视频片段为基于第四映射关系，从目标素材视频中获取的与被选中素材文本对应的视频片段；素材音频片段为基于第五映射关系，从目标素材音频中获取的与被选中素材文本对应的音频片段；待剪辑视频包括待剪辑媒体视频和待剪辑媒体音频。

本公开实施例中，各模块可以实现上述方法实施例提供的视频处理方法，且能达到相同的技术效果，为避免重复，这里不再赘述。

图 7 为本公开实施例提供的一种电子设备的结构示意图，用于对实现本公开实施例中任意视频处理方法的电子设备进行示例性说明，不应理解为对本公开实施例的具体限定。

如图 7 所示，电子设备 700 可以包括处理器（例如中央处理器、图形处理器等）701，其可以根据存储在只读存储器（ROM）702 中的程序或者从存储装置 708 加载到随机访问存储器（RAM）703 中的程序而执行各种适当的动作和处理。在 RAM 703 中，还存储有电子设备 700 操作所需的各种程序和数据。处理器 701、ROM 702 以及 RAM 703 通过总线 704 彼此相连。输入/输出（I/O）接口 705 也连接至总线 704。

通常，以下装置可以连接至 I/O 接口 705：包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 706；包括例如液晶显示器（LCD）、扬声器、振动器等的输出装置 707；包括例如磁带、硬盘等的存储装置 708；以及通信装置 709。通信装置 709 可以允许电子设备 700 与其他设备进行无线或有线通信以交换数据。虽然示出了具有各种装置的电子设备 700，但是应理解的是，并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以

被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在非暂态计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信装置 709 从网络上被下载和安装，或者从存储装置 708 被安装，或者从 ROM 702 被安装。在该计算机程序被处理器 701 执行时，可以执行本公开实施例提供的任意视频处理方法中限定的功能。

需要说明的是，本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、RF（射频）等等，或者上述的任意合适的组合。

在一些实施方式中，客户端、服务端可以利用诸如 HTTP（HyperText Transfer Protocol，超文本传输协议）之类的任何当前已知或未来研发的网络协议进行通信，并且可以与任意形式或介质的数字数据通信（例如，通信网络）互连。通信网络的示例包括局域网（“LAN”），广域网（“WAN”），网际网（例如，互联网）以及端对端网络（例如，ad hoc 端对端网络），以及任何当前已知或未来研发的网络。

上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

上述计算机可读介质承载有一个或者多个程序，当上述一个或者

多个程序被该电子设备执行时，使得该电子设备：展示待处理视频对应的待处理文本数据；其中，待处理视频包括媒体视频和媒体音频，待处理文本数据对应的文本时间码信息，与媒体视频对应的第一时间线信息存在第一映射关系；文本时间码信息与媒体音频对应的第二时间线信息存在第二映射关系；响应于针对待处理文本数据中的目标文本数据的第一触发操作，基于第一映射关系，从媒体视频中确定目标文本数据对应的目标视频片段；基于第二映射关系，从媒体音频中确定目标文本数据对应的目标音频片段； 基于第一触发操作，针对目标文本数据、目标视频片段和目标音频片段进行相应处理。

在本公开实施例中，可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上述程序设计语言包括但不限于面向对象的程序设计语言，诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言，诸如“C”语言或类似的程序设计语言。程序代码可以完全地在计算机上执行、部分地在计算机上执行、作为一个独立的软件包执行、部分在计算机上部分在远程计算机上执行、或者完全在远程计算机或服务端上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络，包括局域网(LAN)或广域网(WAN)，连接到计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。其中，单元的名称在某种情况下并不构成对该单元本身的限定。

本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如，非限制性地，可以使用的示范类型的硬件逻辑部件包括：现场可编程门阵列（FPGA）、专用集成电路（ASIC）、专用标

准产品（ASSP）、片上系统（SOC）、复杂可编程逻辑设备（CPLD）等等。

在本公开的上下文中，计算机可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。计算机可读介质可以是计算机可读信号介质或计算机可读储存介质。计算机可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。计算机可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦除可编程只读存储器（EPROM 或快闪存储器）、光纤、便捷式紧凑盘只读存储器（CD-ROM）、光学储存设备、磁储存设备、或上述内容的任何合适组合。

以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开中所涉及的公开范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述公开构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开中公开的（但不限于）具有类似功能的技术特征进行互相替换而形成的技术方案。

此外，虽然采用特定次序描绘了各操作，但是这不应当理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和并行处理可能是有利的。同样地，虽然在上面论述中包含了若干具体实现细节，但是这些不应当被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实施例中。相反地，在单个实施例的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

尽管已经采用特定于结构特征和/或方法逻辑动作的语言描述了本主题，但是应当理解所附权利要求书中所限定的主题未必局限于上面描述的特定特征或动作。相反，上面所描述的特定特征和动作仅仅是实现权利要求书的示例形式。

权 利 要 求 书

1. 一种视频处理方法，包括：

展示待处理视频对应的待处理文本数据；其中，所述待处理视频包括媒体视频和媒体音频，所述待处理文本数据对应的文本时间码信息，与所述媒体视频对应的第一时间线信息存在第一映射关系；所述文本时间码信息与所述媒体音频对应的第二时间线信息存在第二映射关系；

响应于针对所述待处理文本数据中的目标文本数据的第一触发操作，基于所述第一映射关系，从所述媒体视频中确定所述目标文本数据对应的目标视频片段；

基于所述第二映射关系，从所述媒体音频中确定所述目标文本数据对应的目标音频片段；

基于所述第一触发操作，针对所述目标文本数据、所述目标视频片段和所述目标音频片段进行相应处理。

2. 根据权利要求 1 所述的方法，其中所述待处理视频还包括至少一个包装对象，所述文本时间码信息与每个包装对象对应的时间线信息存在映射关系；

其中，所述每个包装对象的对象类型包括以下至少一项：

字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包、背景音乐。

3. 根据权利要求 2 所述的方法，其中所述第一触发操作包括以下任一项：

删除操作、拖动操作、复制操作、剪切操作；

所述方法还包括：

响应于针对所述待处理文本数据中的目标文本数据的第一触发操作，从所述至少一个包装对象中确定所述目标文本数据对应的目标包装对象；

基于所述第一触发操作，在所述目标包装对象的对象类型为第一类型中的任一种的情况下，对所述目标包装对象进行相应处理；

在所述目标包装对象的对象类型为第二类型中的任一种的情况下，禁止对所述目标包装对象进行相应处理；

其中，所述第一类型包括以下至少一项：字幕文本、画中画、非字幕文本、贴纸、特效、滤镜、素材包；

所述第二类型包括：背景音乐。

4. 根据权利要求 3 所述的方法，其中所述第一类型包括第三类型和第四类型；

所述在所述目标包装对象的对象类型为第一类型中的任一种的情况下，对所述目标包装对象进行相应处理，包括：

在所述目标包装对象的对象类型为第三类型中的任一种的情况下，对所述目标包装对象中的目标包装片段进行相应处理，所述目标包装片段为所述目标包装对象中与所述目标文本数据对应的部分；

或者，

在所述目标包装对象的对象类型为第四类型中的任一种，且所述目标包装对象的起始时间戳在所述目标文本数据对应的时间码范围内的情况下，对所述目标包装对象的整体进行相应处理；

其中，第三类型包括以下至少一项：字幕文本、画中画；

所述第四类型包括以下至少一项：非字幕文本、贴纸、特效、滤镜、素材包。

5. 根据权利要求 4 所述的方法，其中所述方法还包括：

在所述目标包装对象的对象类型为所述第四类型中的任一种，且所述目标包装对象的起始时间戳不在所述目标文本数据对应的时间码范围内的情况下，禁止对所述目标包装对象进行相应处理。

6. 根据权利要求 1 所述的方法，其中所述方法还包括：

响应于针对所述待处理文本数据中的第一文本数据的修改操作，

获取与所述修改操作对应的修改后的第二文本数据；

基于所述第二映射关系，从所述媒体音频中确定与所述第一文本数据对应的第一音频片段；

基于所述第二文本数据和所述第一音频片段的音色信息生成第二音频片段；

将所述媒体音频中的所述第一音频片段替换为所述第二音频片段。

7. 根据权利要求 6 所述的方法，其中所述待处理视频还包括字幕文本，所述文本时间码信息与所述字幕文本对应的第三时间线信息存在第三映射关系；所述方法还包括：

响应于针对所述待处理文本数据中的第一文本数据的修改操作，基于所述第三映射关系，从所述字幕文本中确定与所述第一文本数据对应的第一字幕片段；

基于所述第二文本数据，对所述第一字幕片段进行修改，得到第二字幕片段；

将所述字幕文本中的所述第一字幕片段替换为所述第二字幕片段。

8. 根据权利要求 1 至 7 中任一项所述的方法，其中所述第一触发操作还包括选中操作；所述响应于针对所述待处理文本数据中的目标文本数据的第一触发操作，基于所述第一映射关系，从所述媒体视频中确定所述目标文本数据对应的目标视频片段之前，所述方法还包括：

展示第一时间轴轨道和第二时间轴轨道，所述第一时间轴轨道为所述媒体视频对应的时间轴轨道，所述第二时间轴轨道为所述媒体音频对应的时间轴轨道；

所述基于所述第一触发操作，针对所述目标文本数据、所述目标视频片段和所述目标音频片段进行相应处理，包括：

按照第一显示样式，对所述目标文本数据进行显示；

按照第二显示样式，对所述第一时间轴轨道中与所述目标视频片段对应的部分进行显示；

按照第三显示样式，对第二时间轴轨道中与所述目标音频片段对应的部分进行显示。

9. 根据权利要求 1 至 7 中任一项所述的方法，其中所述方法还包括：

响应于针对所述待处理文本数据中的被选中文本数据的第二触发操作，在所述被选中文本数据对应的区域展示工具菜单，所述工具菜单包括以下至少一类：剪辑控件、添加控件，所述剪辑控件包括以下至少一项：与删除操作对应的删除控件、与复制操作对应的复制控件、与剪切控件对应的剪切控件、与修改操作对应的修改控件，智能长短句控件、搜索控件、替换控件；所述添加控件包括以下至少一项：添加画中画控件、添加非字幕文本控件、添加贴纸控件、添加特效控件、添加音效控件、添加滤镜控件、添加素材包控件、添加背景音乐控件；

响应于对所述工具菜单中的目标控件的第三触发操作，对所述媒体视频、所述待处理文本数据、所述媒体音频中的至少一项执行相应处理。

10. 根据权利要求 1 至 7 中任一项所述的方法，其中所述展示待处理视频对应的待处理文本数据之后，所述方法还包括：

响应于对素材库控件的触发操作，展示多个预设视频素材；

响应于对所述多个预设视频素材中的目标视频素材的第五触发操作，展示文本粗剪浮窗，所述文本粗剪浮窗中包括所述目标视频素材对应的素材文本数据，所述目标视频素材包括目标素材视频和目标素材音频，所述素材文本数据的素材时间码信息与所述目标素材视频的时间线信息存在第四映射关系，所述素材时间码信息与所述目标素材音频的时间线信息存在第五映射关系；

响应于对所述素材文本数据中的被选中素材文本的文本粗剪操

作，将所述被选中素材文本与所述待处理文本数据合成，得到待剪辑文本数据，将素材视频片段与所述媒体视频合成为待剪辑媒体视频，将素材音频片段与所述媒体音频合成为待剪辑媒体音频，以使基于对所述待剪辑文本数据的文本剪辑操作，实现对待剪辑视频的剪辑；

其中，所述素材视频片段为基于所述第四映射关系，从所述目标素材视频中获取的与所述被选中素材文本对应的视频片段；所述素材音频片段为基于所述第五映射关系，从所述目标素材音频中获取的与所述被选中素材文本对应的音频片段；所述待剪辑视频包括所述待剪辑媒体视频和所述待剪辑媒体音频。

11. 一种视频处理装置，包括：

展示模块，用于展示待处理视频对应的待处理文本数据；其中，所述待处理视频包括媒体视频和媒体音频，所述待处理文本数据对应的文本时间码信息，与所述媒体视频对应的第一时间线信息存在第一映射关系；所述文本时间码信息与所述媒体音频对应的第二时间线信息存在第二映射关系；

确定模块，用于响应于针对所述待处理文本数据中的目标文本数据的第一触发操作，基于所述第一映射关系，从所述媒体视频中确定所述目标文本数据对应的目标视频片段；

确定模块，还用于基于所述第二映射关系，从所述媒体音频中确定所述目标文本数据对应的目标音频片段；

处理模块，用于基于所述第一触发操作，针对所述目标文本数据、所述目标视频片段和所述目标音频片段进行相应处理。

12. 一种电子设备，包括：存储器和处理器，存储器用于存储计算机程序；处理器用于在调用计算机程序时执行权利要求 1 至 10 中任一项所述的视频处理方法。

13. 一种计算机可读存储介质，其上存储有计算机程序，计算机程序被处理器执行时实现权利要求 1 至 10 中任一项所述的视频处理方

法。