

说明书

视频生成方法、装置、设备、存储介质和程序产品

本申请要求 2022 年 5 月 10 日递交的，标题为“视频生成方法、装置、设备、存储介质和程序产品”、申请号为 202210508063.2 的中国发明专利申请的优先权。

技术领域

本公开涉及视频处理技术领域，尤其涉及一种视频生成方法、装置、设备、存储介质和程序产品。

背景技术

随着计算机技术和移动通信技术的迅速发展，基于电子设备的各种视频平台得到了普遍应用，极大地丰富了人们的日常生活。越来越多的用户乐于在视频平台上分享自己的视频作品，以供其他用户观看。

在相关技术中，在制作视频时，用户首先需要自行寻找视频中需要的各类素材，然后对素材进行一系列复杂的视频剪辑操作，最终生成一个视频作品。

如果用户缺乏剪辑经验，导致制作视频的时间成本增加，且使得制作的视频质量不高。

发明内容

为了解决上述技术问题，本公开实施例提供了一种视频生成方法、装置、设备、存储介质和程序产品，将获取到的剪辑模板中的剪辑操作直接应用在多媒体数据中，生成视频，也无需用户手动剪辑视频，不但能够降低制作视频的时间成本，还能够提高制作的视频质量。

第一方面，本公开实施例提供一种视频生成方法，包括：

基于接收到的文本数据生成初始多媒体数据；其中，初始多媒体数据包括文本数据的朗读语音与文本数据匹配的视频图像，初始多媒体数据包括至少一个多媒体片段，至少一个多媒体片段分别对应于文本数据划分的至少一个文本片段；至少一个多媒体片段中的目标多媒体片段与至少一个文本片段中的目标文本片段相对应，目标多媒体片

段包括目标视频片段和目标语音片段，目标视频片段包括与目标文本片段匹配的视频图像，目标语音片段包括与目标文本片段匹配的朗读语音；

响应于剪辑模板获取请求，获取目标剪辑模板；

将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据；

基于目标多媒体数据生成目标视频。

第二方面，本公开实施例提供一种视频生成装置，包括：

初始多媒体数据生成模块，用于基于接收到的文本数据生成初始多媒体数据；其中，初始多媒体数据包括文本数据的朗读语音与文本数据匹配的视频图像，初始多媒体数据包括至少一个多媒体片段，至少一个多媒体片段分别对应于文本数据划分的至少一个文本片段；至少一个多媒体片段中的目标多媒体片段与至少一个文本片段中的目标文本片段相对应，目标多媒体片段包括目标视频片段和目标语音片段，目标视频片段包括与目标文本片段匹配的视频图像，目标语音片段包括与目标文本片段匹配的朗读语音；

目标剪辑模板获取模块，用于响应于剪辑模板获取请求，获取目标剪辑模板；

目标多媒体数据生成模块，用于将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据；

目标视频生成模块，用于基于目标多媒体数据生成目标视频。第三方面，本公开实施例提供一种电子设备，电子设备包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序；

当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现如上述第一方面中任一项的视频生成方法。

第四方面，本公开实施例提供一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如上述第一方面中任一项所述的视频生成方法。

第五方面，本公开实施例提供一种计算机程序产品，该计算机程序产品包括计算机程序或指令，该计算机程序或指令被处理器执行时

实现如上述第一方面中任一项所述的视频生成方法。

本公开实施例提供了一种视频生成方法、装置、设备、存储介质和程序产品，方法包括：基于接收到的文本数据生成初始多媒体数据；响应于剪辑模板获取请求，获取目标剪辑模板；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据；基于目标多媒体数据生成目标视频。本公开实施例通过将获取到的剪辑模板中的剪辑操作直接应用多媒体数据中，生成视频，无需用户手动剪辑视频，不但能够降低制作视频的时间成本，还能够提高制作的视频质量。

附图说明

结合附图并参考以下具体实施方式，本公开各实施例的上述和其他特征、优点及方面将变得更加明显。贯穿附图中，相同或相似的附图标记表示相同或相似的元素。应当理解附图是示意性的，原件和元素不一定按照比例绘制。

图1示出了本公开实施例提供的一种视频制作场景的架构图；

图2为本公开实施例中的一种视频生成方法的流程示意图；

图3为本公开实施例中的模板主题控件的触发示意图；

图4为本公开实施例中的模板控件的触发示意图；

图5为本公开实施例中的模板应用提示的示意图；

图6为本公开实施例中的一种视频生成装置的结构示意图；

图7为本公开实施例中的一种电子设备的结构示意图。

具体实施方式

下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例，然而应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为限于这里阐述的实施例，相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是，本公开的附图及实施例仅用于示例性作用，并非用于限制本公开的保护范围。

应当理解，本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行，和/或并行执行。此外，方法实施方式可以包括附加的

步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。

本文使用的术语“包括”及其变形是开放性包括，即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施例”。其他术语的相关定义将在下文描述中给出。

需要注意，本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

需要注意，本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的，本领域技术人员应当理解，除非在上下文另有明确指出，否则应该理解为“一个或多个”。

本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的，而并不是用于对这些消息或信息的范围进行限制。

在对本申请实施例进行详细地解释说明之前，先对本申请实施例的应用场景进行说明。

用户在处理文档时，大多是以文本形式呈现，而用户阅读文本较为费力，因此，可以将文本信息转换成视频，这样，用户可以收听音频以及观看视频画面来明确文章所传递的信息，而不需要费力解读文本，可以降低用户获取信息的难度。或者，由于文本篇幅较长，而用户阅读文本较为费时，用户没有精力逐篇阅读，因此，需要将文章转换成视频，通过视频快速了解文章所传递的信息，再选择自己感兴趣的文章进行仔细阅读。另外，由于视频的展现形式较为多元化，相比于枯燥的文字阅读来说，更容易吸引用户的注意力，用户也更愿意通过这种方式来阅读文章。

相关技术中，需要从文本数据中提取关键词；对于每个关键词，在预先设定的图片库中查找与该关键词相匹配的视频图片；根据排版规则对文本信息和视频图片进行合成，得到目标视频。然而，相关技术中，仅仅是将查找到视频图片与文本数据进行简单的合成，制作的视频质量不高，后续还需要用户手动进行剪辑，如果用户缺乏剪辑经

验，则影响了视频的质量。

本申请实施例中，根据文本数据生成初始多媒体数据后，获取一个目标剪辑模板，将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，从而实现对初始多媒体数据的剪辑处理，无需用户手动剪辑视频，不但能够降低制作视频的时间成本，还能够提高制作的视频质量。图1示出了本公开实施例提供的一种视频制作场景的架构图。

如图1所示，该架构图中可以包括客户端的至少一个电子设备101以及服务端的至少一个服务器102。电子设备101可以通过网络协议如超文本传输安全协议(Hyper Text Transfer Protocol over Secure Socket Layer, HTTPS)与服务器102建立连接并进行信息交互。其中，电子设备101可以包括移动电话、平板电脑、台式计算机、笔记本电脑、车载终端、可穿戴设备、一体机、智能家居设备等具有通信功能的设备，也可以包括虚拟机或者模拟器模拟的设备。服务器102可以包括云服务器或者服务器集群等具有存储及计算功能的设备。

基于上述架构，用户可以在电子设备101上的指定平台内制作视频，指定平台可以为指定应用程序或者指定网站。用户可以在制作好视频后，向指定平台的服务器102发送该视频，服务器102可以接收电子设备101发送的视频，并且存储接收到的视频，以将该视频发送给需要播放该视频的电子设备。

在本公开实施例中，为了降低制作视频的时间成本以及提高制作的视频质量，电子设备101能够接收用户针对初始多媒体数据的剪辑模板获取请求，在电子设备101接收到该剪辑模板获取请求之后，可以获取目标剪辑模板，将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，并基于目标多媒体数据生成目标视频。可见，在目标视频的生成过程中将获取到的目标剪辑模板中的剪辑操作直接应用在初始多媒体数据中，无需用户手动剪辑视频，不但能够降低制作视频的时间成本，还能够提高制作的视频质量。

可选地，基于上述架构，电子设备101还可以在接收到剪辑模板获取请求获取目标剪辑模板，将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，并基于目标多媒体数据生成目标视频，从而在电子设备101本地将目标剪辑模板所指示的剪辑

操作应用于初始多媒体数据，进而生成目标视频，以进一步降低制作视频的时间成本。

可选地，基于上述架构，电子设备 101 还可以在接收到剪辑模板获取请求之后，向服务器 102 发送携带有模板标识的剪辑模板获取请求。服务器 102 可以在接收到电子设备 101 发送的携带有模板标识的剪辑模板获取请求之后，响应于剪辑模板获取请求，获取目标剪辑模板，将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，并基于目标多媒体数据生成目标视频，并向电子设备 101 发送生成的目标视频，从而使电子设备 101 可以请求服务器 102 基于剪辑模板获取请求，获取目标剪辑模板，并将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，生成目标视频，以进一步提高制作的视频质量并且降低电子设备 101 的数据处理量。

例如：电子设备可以是移动终端、固定终端或便携式终端，例如移动手机、站点、单元、设备、多媒体计算机、多媒体平板、互联网节点、通信器、台式计算机、膝上型计算机、笔记本计算机、上网本计算机、平板计算机、个人通信系统(PCS)设备、个人导航设备、个人数字助理(PDA)、音频/视频播放器、数码相机/摄像机、定位设备、电视接收器、无线电广播接收器、电子书设备、游戏设备或者其任意组合，包括这些设备的配件和外设或者其任意组合。

服务器可以是实体服务器，也可以是云服务器，服务器可以是一个服务器，或者服务器集群。

下面将结合附图，对本申请实施例提出的视频生成方法进行详细介绍。

图 2 为本公开实施例中的一种视频生成方法的流程图，本实施例可适用于根据文本信息生成视频的情况，该方法可以由视频生成装置执行，该视频生成装置可以采用软件和/或硬件的方式实现，该视频生成方法可由图 1 中所述的电子设备中。

如图 2 所示，本公开实施例提供的视频生成方法主要包括步骤 S101-S104。

S101、基于接收到的文本数据生成初始多媒体数据。

在本公开的一个实施例中，文本数据可以是用户通过输入装置录

入至电子设备的数据，还可以是其他设备向电子设备发送的数据。

在本公开的一个实施例中，在基于接收到的文本数据生成初始多媒体数据之前，还包括：响应于用户的输入数据操作，接收文本数据。其中，用户的输入数据操作可以包括对文本数据的添加操作，也可以包括对文本数据的录入操作，本实施例中不作具体限制。

在本公开的一个实施例中，初始多媒体数据包括文本数据的朗读语音与文本数据匹配的视频图像，初始多媒体数据包括至少一个多媒体片段，至少一个多媒体片段分别对应于文本数据划分的至少一个文本片段；至少一个多媒体片段中的目标多媒体片段与至少一个文本片段中的目标文本片段相对应，目标多媒体片段包括目标视频片段和目标语音片段，目标视频片段包括与目标文本片段匹配的视频图像，目标语音片段包括与目标文本片段匹配的朗读语音。

在本公开的一个实施方式中，基于接收到的文本数据生成初始多媒体数据，包括：将接收到的文本数据划分的至少一个文本片段，文本片段包括多个目标文本片段。针对每一个目标文本片段，基于目标文本片段在预设图库中查找与目标文本片段对应的视频图像，将视频图像按照预设的动画效果进行处理，得到与目标文本片段对应的目标视频片段。获取目标文本片段匹配的朗读语音，生成目标语音片段。将目标视频片段和目标语音片段进行合成，得到目标多媒体片段。针对每一个目标文本片段，得到多个目标多媒体片段，将多个目标多媒体片段按照目标文本片段的前后顺序进行合成，得到初始多媒体数据。

在本公开的一个实施方式中，视频图像上包括与目标文本片段匹配的字幕文本。

在本公开实施例中，在视频图像上添加与目标文本片段匹配的字幕文本，以方便用户在观看视频的过程中，能够直观的看到与朗读语音对应的字幕，提高用户的观看体验。

S102、响应于剪辑模板获取请求，获取目标剪辑模板。

在本公开的一个实施方式中，响应于剪辑模板获取请求可以是接收到用户对电子设备的操作后，响应于剪辑模板获取请求。还可以是检测到初始多媒体数据生成之后，响应于剪辑模板获取请求。

目标剪辑模板，可以是基于用户对电子设备的操作，选中的剪辑

模板，还可以是基于文本数据中的关键字自动匹配到剪辑模板。

本公开的一个实施例方式中，获取目标剪辑模板，包括：电子设备直接在本机预先存储的模板数据库中获取目标剪辑模板。

在本公开的一个实施方式中，获取目标剪辑模板，包括：电子设备获取目标剪辑模板对应的模板标识，将携带有模板标识的剪辑模板获取请求发送至服务器，服务器响应该携带有模板标识的剪辑模板获取请求，并基于模板标识获取目标剪辑模板，将获取到的目标剪辑模板返回至电子设备。

在本公开的一个实施方式中，如果未能获取目标剪辑模板，则在电子设备的显示界面中展示一个提示弹出框，该提示弹出框用于提示用于目标剪辑模板获取失败。

在本公开的一个实施方式中，响应于剪辑模板获取请求，获取目标剪辑模板，包括：响应于对模板主题控件的触发操作，将触发操作对应的剪辑模板确定为目标剪辑模板；获取目标剪辑模板。

在本公开的一个实施方式中，电子设备的交互界面上显示至少一个模板主题控件，响应于用户对模板主题控件的触发操作，将触发操作对应的剪辑模板确定为目标剪辑模板。

如图 3 所示，响应于用户对模板主题 1 控件的触发操作，将模板主题 1 控件对应的剪辑模板确定为目标剪辑模板。

在本公开实施例中，通过用户的触发操作，选择目标剪辑模板，便于用户选择到自己满意的剪辑模板，提高用户的使用体验。

在本公开的一个实施方式中，响应于对剪辑模板控件的触发操作之前，还包括：显示视频编辑区域，其中，视频编辑区域中包括模板控件；响应于对模板控件的触发操作，显示蒙层区域；在蒙层区域上显示至少一个模板主题控件。

在本公开实施例中，如图 4 所示，生成初始多媒体数据后，在电子设备的显示界面中，显示视频预览区域 10 和视频编辑区域 20，在视频编辑区域 20 中包括多个编辑控件，例如：模板控件、画面控件、文字控件、朗读音色控件和音乐控件。其中，模板控件用于指示用户可以使用现有模板对初始多媒体数据进行编辑。画面控件用于指示用户对初始多媒体数据中的视频图像进行编辑。文字控件用于指示用户对

初始多媒体数据中的字幕文本进行编辑。朗读音色控件用于指示用户对初始多媒体数据中的朗读语音进行编辑。音乐控件用于指示用户对初始多媒体数据中的背景音乐进行编辑。

在本公开的一个实施方式中，如图 4 所示，响应于用户对模板控件的触发操作，显示一个蒙层区域，在蒙层区域中显示多个剪辑模板主题控件。响应于对蒙层区域的左右滑动操作，以左右滑动的效果展示多个剪辑模板主题控件。

在本公开实施例中，响应用户对模板控件的触发操作后，显示多个模板主题控件，使得操作简单易懂，方便用户操作。

S103、将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据。

在本公开的一个实施方式中，目标剪辑模板包括至少一个剪辑操作，该剪辑操作应用于初始多媒体数据上，可以对初始多媒体数据进行剪辑操作。

在本公开的一个实施方式中，如图 5 所示，将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据的过程中，由于对初始多媒体数据的剪辑需要一定的时间，因此，在电子设备的显示界面中展示应用提示框，该应用提示框用于提示用户正在使用剪辑模板中所指示的剪辑操作对初始多媒体视频进行剪辑处理。

在本公开的一个实施方式中，如果目标剪辑模板所指示的剪辑操作成功应用于初始多媒体数据，则显示剪辑模板应用成功的提示消息；如果目标剪辑模板所指示的剪辑操作不成功应用于初始多媒体数据，则显示剪辑模板应用失败的提示消息，并提示用户重新选择剪辑模板。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：视频合成操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，包括：基于视频合成操作将目标剪辑模板中包括的视频片段与初始多媒体数据中包括的多媒体片段进行合成，得到目标多媒体数据。

在本公开的实施例方式中，目标剪辑模板中包括一个或多个视频片段。在目标剪辑模板所指示的剪辑操作包括：视频合成操作的情况下，将目标剪辑模板中包括的一个或多个视频片段与初始多媒体数据

中包括的多媒体片段进行合成，得到目标多媒体数据。

在本公开的实施例方式中，将目标剪辑模板中包括的视频片段添加至多媒体片段的任意两个视频帧之间。上述视频片段合成操作，可以是现有的任意一种视频合成方式，本实施例中不再具体的限定。

在本公开实施例中，通过剪辑模板中的视频合成操作，实现多段视频的合成，避免用户手动合成视频，降低制作视频的时间成本，提高制作的视频质量。

在本公开的一个实施方式中，基于视频合成操作将目标剪辑模板中包括的视频片段与初始多媒体数据中包括的多媒体片段进行合成，得到目标多媒体数据，包括：基于视频合成操作，将目标剪辑模板中包括的视频片段加载至初始多媒体数据中包括的多媒体片段的设定位置，得到目标多媒体数据，其中，设定位置包括：初始多媒体数据第一帧媒体数据之前，和/或，初始多媒体数据最后一帧媒体数据之后。

在本公开实施例中，目标剪辑模板中包括多个视频片段以及各个视频片段对应的添加位置。

在本公开的一个实施方式中，如果目标剪辑模板中包括的视频片段对应的添加位置为片头位置，则将该视频片段添加至初始多媒体数据第一帧媒体数据之前，作为目标视频片头。

在本公开的一个实施方式中，如果目标剪辑模板中包括的视频片段对应的添加位置为片尾位置，则将该视频片段添加至初始多媒体数据最后一帧媒体数据之后，作为目标视频的片头。

在本公开的一个实施方式中，如果文本数据中包括文本主题，将文本主题添加至片头对应的视频片段中的文本主题的位置，并根据目标剪辑模板中包括的文本主题显示效果对文本主题进行编辑并渲染上屏。进一步的，如果文本数据中包括文本作者，将文本作者添加至片头对应的视频片段中的文本作者的位置，并根据目标剪辑模板中包括的文本作者显示效果对文本作者信息进行编辑并渲染上屏。

在本公开的一个实施方式中，如果获取到视频制作者的信息，将视频制作者的信息添加至片尾对应的视频片段中的制作者的位置，并根据目标剪辑模板中包括的视频制作者显示效果对视频制作者的信息进行编辑并渲染上屏。

在本公开实施例中，通过剪辑模板中的视频合成操作，实现添加片头和/或片尾的操作，避免用户手动添加片头或片尾，降低制作视频的时间成本，提高制作的视频质量。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：转场设置操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，包括：基于转场设置操作对初始多媒体数据中包括的多媒体片段添加转场效果，得到目标多媒体数据。

在本公开的一个实施方式中，初始多媒体数据中包括多个与文本数据匹配的视频图像，多个视频图像在切换的过程中，必然涉及到图像转场设置。在相关技术中，用户需要手动设置两个相邻视频图像之间的转场效果，增加视频制作的时间成本。

在本公开的一个实施方式中，转场效果包括如下一个或多个：百叶窗动画效果、切入动画效果、闪烁动画效果、渐变动画效果、十字溶解动画效果、缩放动画效果等等。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：转场设置操作，转场设置操作中包括多个转场效果类型。基于转场设置操作中包括的多个转场效果类型应用于多媒体片段中，使得各个多媒体片段具备各自对应的转场效果。

在本公开的一个实施方式中，如果转场设置操作中包括一个转场效果类型，则将该转场效果类型应用于多媒体片段，使得多媒体片段具备相同的转场效果。

在本公开实施例中，通过剪辑模板中的转场设置操作，为多媒体片段添加转场效果，避免用户手动设置转场效果，降低制作视频的时间成本，提高制作的视频质量。

在本公开的一个实施方式中，将目标剪辑模板所指示的剪辑操作包括：虚拟对象添加操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，包括：基于虚拟对象添加操作将目标剪辑模板中包括的虚拟对象添加至初始多媒体数据的预设位置，得到目标多媒体数据。

在本公开的一个实施方式中，虚拟对象包括：目标视频片段、虚拟贴纸、虚拟物体、虚拟卡片等多种对象。可选的，可以包括：面部

装饰特征、头饰特征、衣物特征和衣物配饰特征等等。

在本公开的一个实施方式中，可以是直接将目标剪辑模板中保存的虚拟对象添加至初始多媒体数据的预设位置。预设位置的具体参数可以保存在目标剪辑模板中，可选的。在保存在目标剪辑模板中设定将闪光效果的贴纸添加在第3幅视频图像上。

在本公开的一个实施方式中，可以根据文本信息中提出的关键字，确定虚拟对象的添加位置。可选的，将虚拟对象添加至关键字对应的视频图像中。

在本公开的实施例中，通过剪辑模板中的虚拟对象添加操作，为多媒体片段添加虚拟对象，避免用户手动添加虚拟对象，降低制作视频的时间成本，提高制作的视频质量。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：背景音频添加操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，包括：基于背景音频添加操作将目标剪辑模板中包括的背景音频与初始多媒体数据中包括的朗读语音进行混合，得到目标多媒体数据。

在本公开的一个实施方式中，目标剪辑模板中包括一个背景音频。基于背景音频添加操作，基于背景音频对应的时间戳和朗读语音对应的时间戳，将背景音频和朗读语音进行混合，得到目标多媒体数据。

在本公开的一个实施方式中，基于朗读语音的播放参数调节背景音频的播放参数，使得两者能够更加融合。

在本公开实施例中，通过剪辑模板中的背景音频的添加操作，为多媒体片段添加背景音乐，避免用户手动添加背景音乐，降低制作视频的时间成本，提高制作的视频质量。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：关键词提取操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，包括：针对至少一个目标文本片段，提取目标文本片段中的关键字；将关键字添加至目标文本片段对应的目标多媒体片段中。

在本公开的一个实施方式中，关键字可以是日期、数字、人物名称、专有名称、地名、植物、动物等等。

在本公开的一个实施方式中，以目标文本片段是“张三于当日向

李四支付现金 20 万元”，该目标文本片段中提取出的关键字为“20 万元”，将关键字“20 万元”添加至该目标文本片段对应的目标多媒体片段中。

在本公开的一个实施方式中，目标剪辑模块中还包括：关键字参数，其中，关键字参数包括：关键字的颜色、字体、添加效果等等。根据关键字参数设置关键字在目标多媒体片段中的显示信息。

在本公开的实施例，通过剪辑模板中的关键字提取操作，为多媒体片段添加关键字，使得用户能够更明确的了解该文本片段的关键信息。

在本公开的一个实施方式中，将关键字添加至目标文本片段对应的目标多媒体片段中，包括：获取与关键字匹配的关键文本信息；将关键字和关键文本信息添加至目标文本片段对应的目标多媒体片段中。

在本公开的实施例，从目标文本片段中提取关键字之后，基于上述关键字获取与关键字匹配的关键信息。例如：关键字是“王五”，与关键字匹配的关键信息为：王五是一名演员，代表作是《电视剧 A》、《电影 B》。此时，将“王五”作为关键字，“演员”，“代表作品《电视剧 A》、《电影 B》”作为关键文本信息，添加至目标多媒体片段中。再如：关键字为“职务侵占罪”，其匹配的关键文本信息为“职务侵占罪，是指公司、企业或者其他单位的人员，利用职务上的便利，将本单位财物非法占为己有，数额较大的行为”此时，将“职务侵占罪”作为关键字，“职务侵占罪，是指公司、企业或者其他单位的人员，利用职务上的便利，将本单位财物非法占为己有，数额较大的行为”作为关键文本信息，添加至目标多媒体片段中。

在本公开的一个实施方式中，可以为关键字、关键文本信息设置不同的显示参数。

在本公开的一个实施方式中，上述与关键字匹配的关键文本信息可以是文本数据中提取出的文本信息，也可以是从互联网或者预设知识库中获取的文本信息。关键文本信息的获取方式本实施例中不再具体限定。

在本公开实施例中，通过关键字提取关键文本信息，将关键字和关键文本信息添加至视频中，使得用户可以快速了解与关键字相关的

知识，辅助用户了解文本数据的内容。

S104、基于目标多媒体数据生成目标视频。

本公开实施例提供了一种视频生成方法、装置、设备、存储介质和程序产品，方法包括：基于接收到的文本数据生成初始多媒体数据；其中，初始多媒体数据包括文本数据的朗读语音与文本数据匹配的视频图像，初始多媒体数据包括至少一个多媒体片段，至少一个多媒体片段分别对应于文本数据划分的至少一个文本片段；至少一个多媒体片段中的目标多媒体片段与至少一个文本片段中的目标文本片段相对应，目标多媒体片段包括目标视频片段和目标语音片段，目标视频片段包括与目标文本片段匹配的视频图像，目标语音片段包括与目标文本片段匹配的朗读语音；响应于剪辑模板获取请求，获取目标剪辑模板；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据；基于目标多媒体数据生成目标视频。本公开实施例通过将获取到的剪辑模板中的剪辑操作直接应用在多媒体数据中，生成视频，无需用户手动剪辑视频，不但能够降低制作视频的时间成本，还能够提高制作的视频质量。

图 6 为本公开实施例中的一种视频生成方法的流程图，本实施例可适用于根据文本信息生成视频的情况，该方法可以由视频生成装置执行，该视频生成装置可以采用软件和/或硬件的方式实现，该视频生成装置可配置于电子设备中。

如图 6 所示，本公开实施例提供的视频生成装置 60 主要包括：初始多媒体数据生成模块 61、目标剪辑模板获取模块 62、目标多媒体数据生成模块 63 和目标视频生成模块 64。

其中，初始多媒体数据生成模块 61，用于基于接收到的文本数据生成初始多媒体数据；其中，初始多媒体数据包括文本数据的朗读语音与文本数据匹配的视频图像，初始多媒体数据包括至少一个多媒体片段，至少一个多媒体片段分别对应于文本数据划分的至少一个文本片段；至少一个多媒体片段中的目标多媒体片段与至少一个文本片段中的目标文本片段相对应，目标多媒体片段包括目标视频片段和目标语音片段，目标视频片段包括与目标文本片段匹配的视频图像，目标语音片段包括与目标文本片段匹配的朗读语音；目标剪辑模板获取模

块 62，用于响应于剪辑模板获取请求，获取目标剪辑模板；目标多媒体数据生成模块 63，用于将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据；目标视频生成模块 64，用于基于目标多媒体数据生成目标视频。

在本公开的一个实施方式中，视频图像上包括与目标文本片段匹配的字幕文本。

在本公开的一个实施方式中，目标剪辑模板获取模块 62，用于响应于剪辑模板获取请求，获取目标剪辑模板时，包括：目标剪辑模板确定单元，用于响应于对模板主题控件的触发操作，将触发操作对应的剪辑模板确定为目标剪辑模板；目标剪辑模板获取单元，用于获取目标剪辑模板。

在本公开的一个实施方式中，目标剪辑模板获取模块 62，还包括：视频编辑区域显示单元，用于响应于对剪辑模板控件的触发操作之前，显示视频编辑区域，其中，视频编辑区域中包括模板控件；蒙层区域显示单元，用于响应于对模板控件的触发操作，显示蒙层区域；在蒙层区域上显示至少一个模板主题控件。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：视频合成操作；目标多媒体数据生成模块 63，具体用于基于视频合成操作将目标剪辑模板中包括的视频片段与初始多媒体数据中包括的多媒体片段进行合成，得到目标多媒体数据。

在本公开的一个实施方式中，目标多媒体数据生成模块 63，具体用于基于视频合成操作，将目标剪辑模板中包括的视频片段加载至初始多媒体数据中包括的多媒体片段的设定位置，得到目标多媒体数据，其中，设定位置包括：初始多媒体数据第一帧媒体数据之前，和/或，初始多媒体数据最后一帧媒体数据之后。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：转场设置操作；目标多媒体数据生成模块 63，具体用于基于转场设置操作对初始多媒体数据中包括的多媒体片段添加转场效果，得到目标多媒体数据。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：虚拟对象添加操作；目标多媒体数据生成模块 63，具体用于基于

虚拟对象添加操作将目标剪辑模板中包括的虚拟对象添加至初始多媒体数据的预设位置，得到目标多媒体数据。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：背景音频添加操作；目标多媒体数据生成模块 63，具体用于基于背景音频添加操作将目标剪辑模板中包括的背景音频与初始多媒体数据中包括的朗读语音进行混合，得到目标多媒体数据。

在本公开的一个实施方式中，目标剪辑模板所指示的剪辑操作包括：关键词提取操作；目标多媒体数据生成模块 63，具体用于针对至少一个目标文本片段，提取目标文本片段中的关键字；将关键字添加至目标文本片段对应的目标多媒体片段中。

在本公开的一个实施方式中，目标多媒体数据生成模块 63，具体用于获取与关键字匹配的关键文本信息；将关键字和关键文本信息添加至目标文本片段对应的目标多媒体片段中。

本公开实施例提供的视频生成装置，可执行本公开方法实施例所提供的视频生成方法中所执行的步骤，具备执行步骤和有益效果此处不再赘述。

图 7 为本公开实施例中的一种电子设备的结构示意图。下面具体参考图 7，其示出了适于用来实现本公开实施例中的电子设备 700 的结构示意图。本公开实施例中的电子设备 700 可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA（个人数字助理）、PAD（平板电脑）、PMP（便携式多媒体播放器）、车载终端（例如车载导航终端）、可穿戴终端设备等等的移动终端以及诸如数字 TV、台式计算机、智能家居设备等等的固定终端。图 7 示出的电子设备仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

如图 7 所示，电子设备 700 可以包括处理装置（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以实现如本公开所述的实施例的图片渲染方法。在 RAM 703 中，还存储有终端设备 700 操作所需的各种程序和数据。处理装置 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 可以允许终端设备 700 与其他设备进行无线或有线通信以交换数据。虽然图 7 示出了具有各种装置的终端设备 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，从而实现如上所述的视频生成方法。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装置 701 执行时，执行本公开实施例的方法中限定的上述功能。

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、服务器可以利用诸如 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终端设备执行时，使得该终端设备：基于接收到的文本数据生成初始多媒体数据；其中，初始多媒体数据包括文本数据的朗读语音与文本数据匹配的视频图像，初始多媒体数据包括至少一个多媒体片段，至少一个多媒体片段分别对应于文本数据划分的至少一个文本片段；至少一个多媒体片段中的目标多媒体片段与至少一个文本片段中的目标文本片段相对应，目标多媒体片段包括目标视频片段和目标语音片段，目标视频片段包括与目标文本片段匹配的视频图像，目标语音片段包括与目标文本片段匹配的朗读语音；响应于剪辑模板获取请求，获取目标剪辑模板；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据；基于目标多媒体数据生成目标视频。

可选的，当上述一个或者多个程序被该终端设备执行时，该终端设备还可以执行上述实施例所述的其他步骤。

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”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个

独立的软件包执行、部分在用户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器上执行。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)——连接到用户计算机，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，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，机器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。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器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。机器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。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，包括：基于接收到的文本数据生成初始多媒体数据；其中，初始多媒体数据包括文本数据的朗读语音与文本数据匹配的视频图像，初始多媒体数据包括至少一个多媒体片段，至少一个多媒体片段分别对应于文本数据划分的至少一个文本片段；至少一个多媒体片段中的目标多媒体片段与至少一个文本片段中的目标文本片段相对应，目标多媒体片段包括目标视频片段和目标语音片段，目标视频片段包括与目标文本片段匹配的视频图像，目标语音片段包括与目标文本片段匹配的朗读语音；响应于剪辑模板获取请求，获取目标剪辑模板；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据；基于目标多媒体数据生成目标视频。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，视频图像上包括与目标文本片段匹配的字幕文本。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，响应于剪辑模板获取请求，获取目标剪辑模板，包括：响应于对模板主题控件的触发操作，将触发操作对应的剪辑模板确定为目标剪辑模板；获取目标剪辑模板。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，响应于对剪辑模板控件的触发操作之前，还包括：显示视频编辑区域，其中，视频编辑区域中包括模板控件；响应于对模板控件的触发操作，显示蒙层区域；在蒙层区域上显示至少一个模板主题控件。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，将目标剪辑模板所指示的剪辑操作包括：视频合成操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，包括：基于视频合成操作将目标剪辑模板中包括的视频片段与初始多媒体数据中包括的多媒体片段进行合成，得到目标多媒体数据。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，基于视频合成操作将目标剪辑模板中包括的视频片段与初始多媒体数据中包括的多媒体片段进行合成，得到目标多媒体数据，包括：基于视频合成操作，将目标剪辑模板中包括的视频片段加载至初始多媒体数据中包括的多媒体片段的设定位置，得到目标多媒体数据，其中，设定位置包括：初始多媒体数据第一帧媒体数据之前，和/或，初始多媒体数据最后一帧媒体数据之后。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，将目标剪辑模板所指示的剪辑操作包括：转场设置操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，包括：基于转场设置操作对初始多媒体数据中包括的多媒体片段添加转场效果，得到目标多媒体数据。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，将目标剪辑模板所指示的剪辑操作包括：虚拟对象添加操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，包括：基于虚拟对象添加操作将目标剪辑模板中包括的虚拟对象添加至初始多媒体数据的预设位置，得到目标多媒体数据。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，目标剪辑模板所指示的剪辑操作包括：背景音频添加操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据，包括：基于背景音频添加操作将目标剪辑模板中包括的背景音频与初始多媒体数据中包括的朗读语音进行混合，得到目标多媒体数据。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，目标剪辑模板所指示的剪辑操作包括：关键词提取操作；将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，包括：针对至少一个目标文本片段，提取目标文本片段中的关键字；将关键字添加至目标文本片段对应的目标多媒体片段中。

根据本公开的一个或多个实施例，本公开提供了一种视频生成方法，其中，将关键字添加至目标文本片段对应的目标多媒体片段中，

包括：获取与关键字匹配的关键文本信息；将关键字和关键文本信息添加至目标文本片段对应的目标多媒体片段中。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，装置包括：初始多媒体数据生成模块，用于基于接收到的文本数据生成初始多媒体数据；其中，初始多媒体数据包括文本数据的朗读语音与文本数据匹配的视频图像，初始多媒体数据包括至少一个多媒体片段，至少一个多媒体片段分别对应于文本数据划分的至少一个文本片段；至少一个多媒体片段中的目标多媒体片段与至少一个文本片段中的目标文本片段相对应，目标多媒体片段包括目标视频片段和目标语音片段，目标视频片段包括与目标文本片段匹配的视频图像，目标语音片段包括与目标文本片段匹配的朗读语音；目标剪辑模板获取模块，用于响应于剪辑模板获取请求，获取目标剪辑模板；目标多媒体数据生成模块，用于将目标剪辑模板所指示的剪辑操作应用于初始多媒体数据，得到目标多媒体数据；目标视频生成模块，用于基于目标多媒体数据生成目标视频。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，视频图像上包括与目标文本片段匹配的字幕文本。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标剪辑模板获取模块，用于响应于剪辑模板获取请求，获取目标剪辑模板时，包括：目标剪辑模板确定单元，用于响应于对模板主题控件的触发操作，将触发操作对应的剪辑模板确定为目标剪辑模板；目标剪辑模板获取单元，用于获取目标剪辑模板。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标剪辑模板获取模块，还包括：视频编辑区域显示单元，用于响应于对剪辑模板控件的触发操作之前，显示视频编辑区域，其中，视频编辑区域中包括模板控件；蒙层区域显示单元，用于响应于对模板控件的触发操作，显示蒙层区域；在蒙层区域上显示至少一个模板主题控件。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标剪辑模板所指示的剪辑操作包括：视频合成操作；目标多媒体数据生成模块，具体用于基于视频合成操作将目标剪辑模板

中包括的视频片段与初始多媒体数据中包括的多媒体片段进行合成，得到目标多媒体数据。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标多媒体数据生成模块，具体用于基于视频合成操作，将目标剪辑模板中包括的视频片段加载至初始多媒体数据中包括的多媒体片段的设定位置，得到目标多媒体数据，其中，设定位置包括：初始多媒体数据第一帧媒体数据之前，和/或，初始多媒体数据最后一帧媒体数据之后。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标剪辑模板所指示的剪辑操作包括：转场设置操作；目标多媒体数据生成模块，具体用于基于转场设置操作对初始多媒体数据中包括的多媒体片段添加转场效果，得到目标多媒体数据。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标剪辑模板所指示的剪辑操作包括：虚拟对象添加操作；目标多媒体数据生成模块，具体用于基于虚拟对象添加操作将目标剪辑模板中包括的虚拟对象添加至初始多媒体数据的预设位置，得到目标多媒体数据。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标剪辑模板所指示的剪辑操作包括：背景音频添加操作；目标多媒体数据生成模块，具体用于基于背景音频添加操作将目标剪辑模板中包括的背景音频与初始多媒体数据中包括的朗读语音进行混合，得到目标多媒体数据。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标剪辑模板所指示的剪辑操作包括：关键词提取操作；目标多媒体数据生成模块，具体用于针对至少一个目标文本片段，提取目标文本片段中的关键字；将关键字添加至目标文本片段对应的目标多媒体片段中。

根据本公开的一个或多个实施例，本公开提供了一种视频生成装置，其中，目标多媒体数据生成模块，具体用于获取与关键字匹配的关键文本信息；将关键字和关键文本信息添加至目标文本片段对应的目标多媒体片段中。

根据本公开的一个或多个实施例，本公开提供了一种电子设备，包括：

一个或多个处理器；

存储器，用于存储一个或多个程序；

当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现如本公开提供的任一所述的视频生成方法。

根据本公开的一个或多个实施例，本公开提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如本公开提供的任一所述的视频生成方法。

本公开实施例还提供了一种计算机程序产品，该计算机程序产品包括计算机程序或指令，该计算机程序或指令被处理器执行时实现如上所述的视频生成方法。

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. 一种视频生成方法，其特征在于，包括：

基于接收到的文本数据生成初始多媒体数据；其中，所述初始多媒体数据包括所述文本数据的朗读语音与所述文本数据匹配的视频图像，所述初始多媒体数据包括至少一个多媒体片段，所述至少一个多媒体片段分别对应于所述文本数据划分的至少一个文本片段；所述至少一个多媒体片段中的目标多媒体片段与所述至少一个文本片段中的目标文本片段相对应，所述目标多媒体片段包括目标视频片段和目标语音片段，所述目标视频片段包括与所述目标文本片段匹配的视频图像，所述目标语音片段包括与所述目标文本片段匹配的朗读语音；

响应于剪辑模板获取请求，获取目标剪辑模板；

将所述目标剪辑模板所指示的剪辑操作应用于所述初始多媒体数据，得到目标多媒体数据；

基于所述目标多媒体数据生成目标视频。

2. 根据权利要求 1 所述的方法，其特征在于，所述视频图像上包括与所述目标文本片段匹配的字幕文本。

3. 根据权利要求 1 所述的方法，其特征在于，响应于剪辑模板获取请求，获取目标剪辑模板，包括：

响应于对模板主题控件的触发操作，将所述触发操作对应的剪辑模板确定为目标剪辑模板；

获取所述目标剪辑模板。

4. 根据权利要求 3 所述的方法，其特征在于，响应于对剪辑模板控件的触发操作之前，还包括：

显示视频编辑区域，其中，所述视频编辑区域中包括模板控件；

响应于对所述模板控件的触发操作，显示蒙层区域；

在所述蒙层区域上显示至少一个模板主题控件。

5. 根据权利要求 1 所述的方法，其特征在于，所述目标剪辑模板所指示的剪辑操作包括：视频合成操作；

将所述目标剪辑模板所指示的剪辑操作应用于所述初始多媒体数据，得到目标多媒体数据，包括：

基于所述视频合成操作将所述目标剪辑模板中包括的视频片段与所述初始多媒体数据中包括的多媒体片段进行合成，得到目标多媒体数据。

6. 根据权利要求 5 所述的方法，其特征在于，基于所述视频合成操作将所述目标剪辑模板中包括的视频片段与所述初始多媒体数据中包括的多媒体片段进行合成，得到目标多媒体数据，包括：

基于所述视频合成操作，将所述目标剪辑模板中包括的视频片段加载至所述初始多媒体数据中包括的多媒体片段的设定位置，得到目标多媒体数据，其中，所述设定位置包括：所述初始多媒体数据第一帧媒体数据之前，和/或，所述初始多媒体数据最后一帧媒体数据之后。

7. 根据权利要求 1 所述的方法，其特征在于，所述目标剪辑模板所指示的剪辑操作包括：转场设置操作；

将所述目标剪辑模板所指示的剪辑操作应用于所述初始多媒体数据，得到目标多媒体数据，包括：

基于所述转场设置操作对所述初始多媒体数据中包括的多媒体片段添加转场效果，得到目标多媒体数据。

8. 根据权利要求 1 所述的方法，其特征在于，所述目标剪辑模板所指示的剪辑操作包括：虚拟对象添加操作；

将所述目标剪辑模板所指示的剪辑操作应用于所述初始多媒体数据，得到目标多媒体数据，包括：

基于所述虚拟对象添加操作将所述目标剪辑模板中包括的虚拟对象添加至所述初始多媒体数据的预设位置，得到目标多媒体数据。

9. 根据权利要求 1 所述的方法，其特征在于，所述目标剪辑模板所指示的剪辑操作包括：背景音频添加操作；

将所述目标剪辑模板所指示的剪辑操作应用于所述初始多媒体数据，得到目标多媒体数据，包括：

基于所述背景音频添加操作将所述目标剪辑模板中包括的背景音频与所述初始多媒体数据中包括的朗读语音进行混合，得到目标多媒体数据。

10. 根据权利要求 1 所述的方法，其特征在于，所述目标剪辑模板所指示的剪辑操作包括：关键词提取操作；

将所述目标剪辑模板所指示的剪辑操作应用于所述初始多媒体数据，包括：

针对至少一个目标文本片段，提取所述目标文本片段中的关键字；

将所述关键字添加至所述目标文本片段对应的目标多媒体片段中。

11. 根据权利要求 10 所述的方法，其特征在于，将所述关键字添加至所述目标文本片段对应的目标多媒体片段中，包括：

获取与所述关键字匹配的关键文本信息；

将所述关键字和所述关键文本信息添加至所述目标文本片段对应的目标多媒体片段中。

12. 一种视频生成装置，其特征在于，包括：

初始多媒体数据生成模块，用于基于接收到的文本数据生成初始多媒体数据；其中，所述初始多媒体数据包括所述文本数据的朗读语音与所述文本数据匹配的视频图像，所述初始多媒体数据包括至少一个多媒体片段，所述至少一个多媒体片段分别对应于所述文本数据划分的至少一个文本片段；所述至少一个多媒体片段中的目标多媒体片段与所述至少一个文本片段中的目标文本片段相对应，所述目标多媒体片段包括目标视频片段和目标语音片段，所述目标视频片段包括与所述目标文本片段匹配的视频图像，所述目标语音片段包括与所述目标文本片段匹配的朗读语音；

目标剪辑模板获取模块，用于响应于剪辑模板获取请求，获取目标剪辑模板；

目标多媒体数据生成模块，用于将所述目标剪辑模板所指示的剪辑操作应用于所述初始多媒体数据，得到目标多媒体数据；

目标视频生成模块，用于基于所述目标多媒体数据生成目标视频。

13. 一种电子设备，其特征在于，所述电子设备包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-11 中任一项所述的方法。

14. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，该程序被处理器执行时实现如权利要求 1-11 中任一项所述的方法。

15. 一种计算机程序产品，该计算机程序产品包括计算机程序或指令，该计算机程序或指令被处理器执行时实现如权利要求 1-11 中任一项所述的方法。