

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2025 年 6 月 26 日 (26.06.2025)



(10) 国际公布号
WO 2025/130705 A1

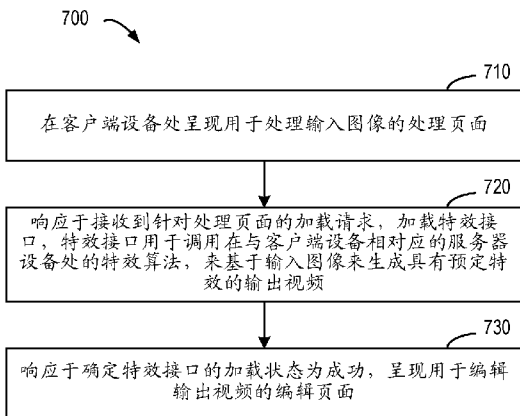
- (51) 国际专利分类号:
H04N 21/472 (2011.01)
- (21) 国际申请号: PCT/CN2024/138191
- (22) 国际申请日: 2024 年 12 月 10 日 (10.12.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202311785540.0 2023年12月22日 (22.12.2023) CN
- (71) 申请人: 北京字跳网络技术有限公司
(**BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区紫金数码园 4 号楼 2 层 0207 100190 (CN)。脸萌有限公司 (**LEMON INC.**) [GB/GB]; 英国英属开曼群

岛大开曼西湾路 802 号木槿道大楼邮政信箱 31119 KY1-1205 (GB)。

- (72) 发明人: 程景亮 (**CHENG, Jingliang**); 中国北京市朝阳区七圣中街 12 号院融中心 B1 小邮局 100028 (CN)。景运鹏 (**JING, Yunpeng**); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室 90066 (US)。厉维凯 (**LI, Weikai**); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室 90066 (US)。马佳欣 (**MA, Jiaxin**); 中国北京市朝阳区七圣中街 12 号院融中心 B1 小邮局 100028 (CN)。陈静洁 (**CHEN, Jing Jie**); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室 90066 (US)。唐萱 (**TANG, Xuan**); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室 90066 (US)。彭浩翔 (**PENG, Haoxiang**); 中国北京市朝阳区七圣中街 12 号院融中心 B1 小邮局

(54) **Title:** METHOD AND APPARATUS FOR GENERATING VIDEO, AND DEVICE AND MEDIUM

(54) 发明名称: 用于生成视频的方法、装置、设备和介质



- 710 Present at a client device a processing page for processing an input image
- 720 In response to a loading request for the processing page being received, load a special effect interface, wherein the special effect interface is used for calling a special effect algorithm at a server device corresponding to the client device to generate an output video having a predetermined special effect on the basis of the input image
- 730 In response to determining that the loading state of the special effect interface is successful, present an editing page for editing the output video

图7

(57) **Abstract:** Provided are a method and apparatus for generating a video, and a device and a medium. The method comprises: presenting at a client device a processing page for processing an input image; in response to a loading request for the processing page being received, loading a special effect interface, wherein the special effect interface is used for calling a special effect algorithm at a server device corresponding to the client device to generate an output video having a predetermined special effect on the basis of the input image; and in response to determining that the loading state of the special effect interface is successful, presenting an editing page for editing the output video. By using the exemplary implementation of the present disclosure, the special effect interface having a powerful video editing capability can be pre-loaded from the server device, and then richer visual effects can be obtained at the client device.

(57) 摘要: 提供了用于生成视频的方法、装置、设备和介质。在一种方法中, 在客户端设备处呈现用于处理输入图像的处理页面。响应于接收到针对处理页面的加载请求, 加载特效接口, 特效接口用于调用在与客户端设备相对应的服务器设备处的特效算法, 来基于输入图像来生成具有预定特效的输出视频。响应于确定特效接口的加载状态为成功, 呈现用于编辑输出视频的编辑页面。利用本公开的示例性实现方式, 可以从服务器设备处预先加载具有强大视频编辑能力的特效接口, 进而在客户端设备处获得更为丰富的视觉效果。

100028 (CN)。高国栋(GAO, Guodong); 中国北京市朝阳区七圣中街12号院融中心B1小邮局 100028 (CN)。张远铭(ZHANG, Yuanming); 中国北京市朝阳区七圣中街12号院融中心B1小邮局 100028 (CN)。王首尧(WANG, Shouyao); 中国北京市朝阳区七圣中街12号院融中心B1小邮局 100028 (CN)。李爽(LI, Shuang); 美国加利福尼亚州卡尔弗城布里斯托大道5800号 90230 (US)。邱韦融(CHIOU, Wei Zon); 美国加利福尼亚州洛杉矶西杰弗逊大道12655号六层第137室 90066 (US)。

(74) 代理人:北京世辉律师事务所(SHIHUI PARTNERS); 中国北京市朝阳区建国门外大街2号北京银泰中心C座42层 100022 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

用于生成视频的方法、装置、设备和介质

5 本申请要求 2023 年 12 月 22 日递交的、标题为“用于生成视频的方法、装置、
设备和介质”、申请号为 202311785540.0 的中国发明专利申请的优先权，该申请的全
部内容通过引用结合在本申请中。

技术领域

10 本公开的示例性实现方式总体涉及视频处理，特别地涉及用于生成视频的方法、
装置、设备和计算机可读存储介质。

背景技术

15 计算机视觉技术已经得到了飞速发展，目前已经开发出了大量媒体编辑工具。例
如，用户可以拍摄照片和/或视频，并且可以使用多种特效接口来在客户端设备处执行
编辑过程，以便生成新的图像和/或视频。然而，客户端设备的处理能力较为有限，因
而期望调用服务器设备的强大处理能力来生成具有更好视觉效果图像和/或视频。

发明内容

20 在本公开的第一方面，提供了一种用于生成视频的方法。在该方法中，在客户端
设备处呈现用于处理输入图像的处理页面。响应于接收到针对处理页面的加载请求，
加载特效接口，特效接口用于调用在与客户端设备相对应的服务器设备处的特效算法，
来基于输入图像来生成具有预定特效的输出视频。响应于确定特效接口的加载状态为
成功，呈现用于编辑输出视频的编辑页面。

25 在本公开的第二方面，提供了一种用于生成视频的装置。该装置包括：第一呈现
模块，被配置用于在客户端设备处呈现用于处理输入图像的处理页面；加载模块，被
配置用于响应于接收到针对处理页面的加载请求，加载特效接口，特效接口用于调用
在与客户端设备相对应的服务器设备处的特效算法，来基于输入图像来生成具有预定
特效的输出视频；以及第二呈现模块，被配置用于响应于确定特效接口的加载状态为
成功，呈现用于编辑输出视频的编辑页面。

30 在本公开的第三方面，提供了一种电子设备。该电子设备包括：至少一个处理单
元；以及至少一个存储器，至少一个存储器被耦合到至少一个处理单元并且存储用于
由至少一个处理单元执行的指令，指令在由至少一个处理单元执行时使电子设备执行
根据本公开第一方面的方法。

在本公开的第四方面，提供了一种计算机可读存储介质，其上存储有计算机程序，计算机程序在被处理器执行时使处理器实现根据本公开第一方面的方法。

应当理解，本内容部分中所描述的内容并非旨在限定本公开的实现方式的关键特征或重要特征，也不用于限制本公开的范围。本公开的其它特征将通过以下的描述而变得容易理解。

附图说明

在下文中，结合附图并参考以下详细说明，本公开各实现方式的上述和其他特征、优点及方面将变得更加明显。在附图中，相同或相似的附图标注表示相同或相似的元素，其中：

图 1 示出了根据本公开的一个示例性实现方式的应用环境的框图；

图 2 示出了根据本公开的一些实现方式的用于生成视频的框图；

图 3 示出了根据本公开的一些实现方式的用于加载特效接口的过程的框图；

图 4 示出了根据本公开的一些实现方式的编辑页面的框图；

图 5 示出了根据本公开的一些实现方式的编辑页面的框图；

图 6 示出了根据本公开的一些实现方式的用于生成视频的交互过程的轨道图；

图 7 示出了根据本公开的一些实现方式的用于生成视频的方法的流程图；

图 8 示出了根据本公开的一些实现方式的用于生成视频的装置的框图；以及

图 9 示出了能够实施本公开的多个实现方式的设备的框图。

具体实施方式

下面将参照附图更详细地描述本公开的实现方式。虽然附图中示出了本公开的某些实现方式，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现方式，相反，提供这些实现方式是为了更加透彻和完整地理解本公开。应当理解的是，本公开的附图及实现方式仅用于示例性作用，并非用于限制本公开的保护范围。

在本公开的实现方式的描述中，术语“包括”及其类似用语应当理解为开放性包含，即“包括但不限于”。术语“基于”应当理解为“至少部分地基于”。术语“一个实现方式”或“该实现方式”应当理解为“至少一个实现方式”。术语“一些实现方式”应当理解为“至少一些实现方式”。下文还可能包括其他明确的和隐含的定义。如本文中所使用的，术语“模型”可以表示各个数据之间的关联关系。例如，可以基于目前已知的和/或将在未来开发的多种技术方案来获取上述关联关系。

可以理解的是，本技术方案所涉及的数据（包括但不限于数据本身、数据的获取

或使用)应当遵循相应法律法规及相关规定的要求。

可以理解的是,在使用本公开各实施例公开的技术方案之前,均应当根据相关法律法规通过适当的方式对本公开所涉及个人信息类型、使用范围、使用场景等告知用户并获得用户的授权。

5 例如,在响应于接收到用户的主动请求时,向用户发送提示信息,以明确地提示用户,其请求执行的操作将需要获取和使用到用户的个人信息。从而,使得用户可以根据提示信息来自主地选择是否向执行本公开技术方案的操作的电子设备、应用程序、服务器或存储介质等软件或硬件提供个人信息。

10 作为一种可选的但非限制性的实现方式,响应于接收到用户的主动请求,向用户发送提示信息的方式,例如可以是弹出窗口的方式,弹出窗口中可以以文字的方式呈现提示信息。此外,弹出窗口中还可以承载供用户选择“同意”或“不同意”向电子设备提供个人信息的选择控件。

可以理解的是,上述通知和获取用户授权过程仅是示意性的,不对本公开的实现方式构成限定,其他满足相关法律法规的方式也可应用于本公开的实现方式中。

15 在此使用的术语“响应于”表示相应的事件发生或者条件得以满足的状态。将会理解,响应于该事件或者条件而被执行的后续动作的执行时机,与该事件发生或者条件成立的时间,二者之间未必是强关联的。例如,在某些情况下,后续动作可在事件发生或者条件成立时立即被执行;而在另一些情况下,后续动作可在事件发生或者条件成立后经过一段时间才被执行。

20

示例环境

25 计算机视觉技术已经得到了飞速发展,目前已经开发出了大量媒体编辑工具。参见图1描述根据本公开的一个示例实现方式的应用环境,该图1示出了根据本公开的一个示例性实现方式的应用环境的框图100。可以在客户端设备处运行媒体编辑工具,该工具例如可以提供如图1所示的页面110。用户可以利用该工具来处理输入图像120。在此,输入图像120例如可以是拍摄的图像或者以其他方式从本地或者远程获取的图像。

30 媒体编辑工具可以提供各种特效控件(也称为特效道具),例如,用户可以使用本地特效控件130来在本地处理输入图像,以便生成新的图像和/或视频。然而,客户端设备的处理能力较为有限,因而期望调用服务器设备的强大处理能力来生成具有更好视觉效果图像和/或视频。

生成视频的概要

为了至少部分地解决现有技术中的不足，根据本公开的一个示例性实现方式，提出了一种用于生成视频的方法。参见图 2 描述根据本公开的一个示例性实现方式的概要，图 2 示出了根据本公开的一些实现方式的用于生成视频的框图 200。如图 2 所示，可以在客户端设备处呈现用于处理输入图像 220 的处理页面 210。处理页面 210 可以包括一个或者多个特效控件，例如，可以提供本地特效控件 130 来以本地方式处理输入图像 220。备选地和/或附加地，可以提供远程特效控件 230 来加载特效接口，从而调用位于服务器设备处的特效算法来处理输入图像 220。

在本公开的上下文中，特效接口可以由多种编程语言来实现，并且表示在客户端设备处实现的过程和/或函数，以便用于作为客户端设备和服务器设备之间的中介，来在客户端设备处调用由服务器设备提供的特效算法。具体地，特效接口可以包括特效工具、特效道具、或者用于调用服务器设备处的特效算法的其他形式。根据本公开的一个示例实现方式，在接收到针对处理页面的加载请求的情况下，可以在客户端设备处加载特效接口。具体地，用户可以点击远程特效控件 230 来提交加载请求，此时所加载的特效接口可以从服务器设备 240 处调用特效算法 242。该特效算法 242 是由加载请求指定的、并且用于基于输入图像 220 来生成具有预定特效的输出视频。

应当理解，尽管图 2 仅仅示意性示出了一个远程特效控件 230，处理页面 210 可以包括一个或者多个远程特效控件；备选地和/或附加地，在用户点击远程特效控件 230 之后，还可以弹出包括多个特效控件的面板以使用户选择。进一步，在确定特效接口的加载状态为成功的情况下，可以呈现用于编辑输出视频的编辑页面 250。

此时，用户可以在编辑页面 250 中执行浏览和各种功能编辑操作。例如，用户可以拖动视频的进度条控件 254 来改变当前所呈现的图像帧。在编辑页面 250 中可以支持用户呈现任意图像帧的切换逻辑。也即，无论用户切换至任何图像帧，所呈现的效果都是正确的。以此方式，可以确保用户在客户端处能够访问由服务器处的远程特效算法产生的视觉效果，由此向用户提供更为丰富的交互体验和更为优秀的视觉效果。

生成视频的详细过程

已经参见图 2 描述了根据本公开的一个示例实现方式的概要，在下文中，将提供有关生成视频的更多细节。根据本公开的一个示例实现方式，编辑页面 250 可以包括：用于呈现输出视频的预览区域 252、以及用于管理输出视频的播放的管理控件。具体地，如图 2 所示，预览区域 252 可以提供输出视频的视觉效果，并且管理控件例如可以包括进度条控件 254、以及用于控制启动播放和暂停播放的切换控件 256。用户可以控制在预览区域 252 中呈现哪个图像帧。以此方式，可以便于用户逐帧查看视觉效果是否符合预期，并且逐帧编辑输出视频。

根据本公开的一个示例实现方式，在加载特效接口的过程中，可以呈现指示特效接口的加载进度的加载页面。参见图 3 描述更多细节，该图 3 示出了根据本公开的一些实现方式的用于加载特效接口的过程的框图 300。如图 3 所示，在加载特效接口的过程中可以呈现编辑页面 310，此时预览区域 252 可以呈现输入图像，并且可以呈现加载页面 320，加载页面 320 例如可以指示加载过程的进度。可以以进度条和/或百分比等方式来向用户呈现已经加载的比例，以此方式，可以便于用户预测等待时间进而合理安排后续的操作。

根据本公开的一个示例实现方式，在确定特效接口的加载状态为成功的情况下，可以呈现指示加载成功的提示信息。类似地，在确定特效接口的加载状态为失败的情况下，呈现指示加载失败的提示信息。以此方式，可以明确地提醒用户当前的加载状态，从而便于用户确定后续的操作步骤。

根据本公开的一个示例实现方式，在加载特效接口的过程中，可以在客户端设备处创建缓存区域。该缓存区域可以用于缓存来自特效算法的多个算法参数。例如，在客户端设备处，可以由媒体编辑工具的客户端来创建缓存目录，该目录可以位于客户端的指定位置处：“***\cache”或者其他位置。进一步，客户端可以通知本地的软件开发工具（SDK）来加载特效接口，继而，SDK 将会执行加载特效接口的具体过程。进一步，客户端可以设置算法的缓存路径，也即上文所创建的缓存目录的路径。进一步，客户端可以执行预处理阶段的处理过程，也即，可以从服务器设备处逐一获取用于生成输出视频中的每个图像帧的多个算法参数。

具体地，可以向缓存区域中写入多个算法参数，可以利用多个算法参数来生成输出视频中的多个图像帧。具体地，可以利用已经加载的特效接口来从服务器设备处获取算法参数。可以从服务器设备处下载各个算法参数。应当理解，在此的算法参数是指概括的概念，算法参数可以包括一组参数（例如，一个或者多个）。应当理解，对于每个特效接口而言，输出视频的长度是固定的，假设输出视频包括 N 个图像帧，则可以从服务器设备下载所需的算法参数，进而基于各个图像帧的在视频中的位置生成相应的图像帧。

根据本公开的一个示例实现方式，在向缓存区域中写入多个算法参数的过程中，可以经由特效接口来从服务器设备获取多个算法参数中的目标算法参数，进而向缓存区域中写入目标算法参数。在需要调用时，只需从客户端设备处本地的缓存目录中查找，而不必从服务器处下载。以此方式，可以减少后续等待时间，进而提高视频编辑过程的处理速度。

应当理解，在访问服务器来获取某个目标算法参数之前，可以首先查看在缓存目录中是否存在已经下载的算法参数，并且仅在缓存目录中不存在该算法参数的情况下，

才访问服务器设备来获取期望的算法参数；否则，可以直接使用被缓存的算法参数。

根据本公开的一个示例实现方式，在获取目标算法参数期间，仅在确定在缓存区域中不存在目标算法参数的情况下，才从服务器设备获取目标算法参数。利用本公开的示例实现方式，缓存目录可以降低访问服务器设备的频率，进而以更为高效的方式获取各个算法参数。

根据本公开的一个示例实现方式，在已经将特效接口成功加载到本地的情况下，可以在编辑页面中编辑输出视频中的各个图像帧。在接收到针对管理控件的管理请求的情况下，可以确定将被呈现在预览区域中的目标图像帧在输出视频中的目标位置。进而利用已经加载到客户端设备本地的特效接口所对应的多个算法参数，生成对应于目标位置的目标图像帧，进而在预览区域 252 中呈现目标图像帧。

参见图 4 描述更多细节，该图 4 示出了根据本公开的一些实现方式的编辑页面的框图 400。在特效接口加载成功的初始阶段，可以呈现如图 4 所示的编辑页面 410。此时，预览区域 252 呈现最初的输入图像，用户可以与管理控件进行交互。具体地，用户可以点击切换控件 256 来开始播放，或者可以调整进度条控件 254 来切换至某个图像帧。利用本公开的示例实现方式，用户可以使用管理控件来方便地切换至任意图像帧，进而执行查看和/或编辑操作。

应当理解，在编辑页面 410 的框 420 处，可以呈现一个或者多个编辑工具。用户可以使用这些编辑工具来调整输出视频，例如，可以向输出视频中添加背景音乐、添加文字，等等。备选地和/或附加地，可以在框 420 处呈现一个或者多个其他特效接口。用户可以选择其他特效接口，此时，如果被选择的特效接口为服务器端的特效接口，则可以从服务器设备处加载该特效接口，进而向当前的输入图像应用该特效接口，进而生成新的输出视频。

根据本公开的一个示例实现方式，假设用户选择播放某个目标图像帧，则可以基于该目标图像帧的目标位置来生成目标图像帧。参见图 5 描述更多细节，该图 5 示出了根据本公开的一些实现方式的编辑页面的框图 500。在编辑页面 510 中，用户可以将进度条控件 254 拖动至图 5 所示的位置（例如，第 i 个图像帧），可以基于输入图像、位置 i 、以及缓存区域中的多个算法参数来生成目标图像帧。

如图 5 所示，可以利用从本地缓存区域中获取的多个算法参数来生成第 i 帧处的图像。预览区域 252 将呈现如图 5 所示的效果，此时在人物周围出现闪电效果。备选地和/或附加地，用户可以将进度条控件 254 进一步调整至第 j 帧的位置。此时，在预览区域中可以呈现图像帧 520 的效果（例如，黑白翻转效果）。利用本公开的示例实现方式，不必从服务器设备处下载期望的算法参数，而是可以直接从本地获取算法参数进而生成目标图像帧。以此方式，可以提高图像生成效率并且降低用户等待时间。

根据本公开的一个示例实现方式，用户可以保存由特效接口生成的输出视频；备选地和/或附加地，用户可以丢弃该输出视频。在接收到用于退出编辑页面的退出请求的情况下，特效接口的生命周期结束因而可以从客户端设备处移除特效接口。以此方式，可以避免不再使用的特效接口占用客户端设备的各种资源。

5 根据本公开的一个示例实现方式，用户可以切换当前正在处理的输入图像。在接收到用于切换输入图像的切换请求的情况下，可以清除缓存区域。应当理解，缓存区域中的算法参数专用于特定的输入图像，因而当输入图像被改变的情况下，缓存区域中的算法参数不再有效，因而应当清除这些参数以避免占用客户端设备的本地存储空间。

10 根据本公开的一个示例实现方式，仅在确定用户选择的特效算法为远程类型的情况下，才从服务器设备加载特效接口。以此方式，可以在媒体编辑工具处维护多种类型的特效接口，从而便于向用户提供更为丰富的视觉处理效果。

已经描述了利用特效接口生成视频的前端操作页面，在下文中，参见图 6 描述在生成视频过程中的各个模块之间的后台交互过程。图 6 示出了根据本公开的一些实现方式的用于生成视频的交互过程的轨道图 600。如图 6 所示，客户端 630 可以表示被安装在客户端设备处的媒体编辑工具的客户端，软件开发工具包 632 表示客户端 630 能够调用的各种功能，特效接口 634 表示用于借助于服务器处的特效算法来将输入图像转换至输出视频的转换工具，缓存目标 636 表示在客户端设备处的预定位置的目录。

15 如图 6 所示，整个交互过程可以划分为预处理阶段 640 和工具使用阶段 642。预处理阶段 640 的主要功能为加载工具，通过工具访问服务端获得算法参数或者读取已经缓存在本地的算法参数，并且发送消息给客户端，使客户端了解是否获得工具正常运行所需要的算法参数。在工具使用阶段 642，用户可以正常使用已经加载的特效接口，可以将播放进度条移动到任意一帧并且查看效果。

25 在初始阶段，用户 638 可以加载 601 特效接口。客户端 630 可以基于该加载操作来创建 602 本地的缓存目录。继而，客户端 630 可以通知 603 软件开发工具包 632 来执行加载过程，并且软件开发工具包 632 可以加载 604 特效接口。客户端 630 可以设置 605 用于缓存算法参数的缓存目录，并且执行 606 预处理。

在预处理过程中，可以访问 610 服务器设备，并且获取相应的多个算法参数。可以将获取的算法参数写入 611 缓存目录。在写入成功的情况下，缓存目录 636 可以向客户端 630 返回 612 成功消息。备选地和/或附加地，可以存在访问失败的情况，此时特效接口 634 将返回 613 失败消息，并且可以向用户显示 614 失败消息。

应当理解，特效接口 634 仅在缓存目录中不存在期望的算法参数的情况下，才访问服务器设备。如果在缓存目录中存在期望的算法参数，则特效接口 634 可以直接读

取 615 缓存目录。缓存目录 636 可以向特效接口 634 返回已经存储的算法参数，并且特效接口 634 将向客户端 630 返回 617 成功消息。至此，预处理阶段 640 结束，并且在缓存目录中已经包括特效处理所需的全部算法参数。

在用户使用阶段 642 中，用户可以将进度条移动 620 至目标视频中的任何位置。

5 此时，客户端 630 可以查找 621 相应的视频帧（例如，第 i 帧）。继而，可以直接从缓存目录中获取所需的算法参数，并基于输入图像和位置 i 来生成第 i 个图像帧。如果用户不再使用 622 加载的特效接口，客户端 630 可以通知 623 软件开发工具包 632 来执行卸载，继而软件开发工具包 632 可以卸载 624 特效接口 634。备选地和/或附加地，如果用户更换 624 输入图像，则客户端 630 可以删除 626 缓存目录。

10 利用本公开的示例实现方式，在编辑页场景下，预处理阶段可以预先获取服务端处的算法参数，从而确保特效接口的高速运行。进一步，缓存目录可以减少对服务器设备的访问次数。缓存目录中包括同样输入的算法参数，因而可以通过重用算法参数来降低用户等待时间。

15 示例过程

图 7 示出了根据本公开的一些实现方式的用于生成视频的方法 700 的流程图。在框 710 处，在客户端设备处呈现用于处理输入图像的处理页面。在框 720 处，响应于接收到针对所述处理页面的加载请求，加载特效接口，所述特效接口用于调用在与所述客户端设备相对应的服务器设备处的特效算法，来基于所述输入图像来生成具有预
20 定特效的输出视频。在框 730 处，响应于确定特效接口的加载状态为成功，呈现用于编辑输出视频的编辑页面。

根据本公开的一个示例实现方式，编辑页面包括：用于呈现输出视频的预览区域、以及用于管理输出视频的播放的管理控件。

根据本公开的一个示例实现方式，该方法进一步包括：响应于接收到针对管理控
25 件的管理请求，确定将被呈现在预览区域中的目标图像帧在输出视频中的目标位置；利用特效接口，生成对应于目标位置的目标图像帧；以及在预览区域中呈现目标图像帧。

根据本公开的一个示例实现方式，该方法进一步包括：呈现指示特效接口的加载进度的加载页面。

30 根据本公开的一个示例实现方式，加载特效接口包括：在客户端设备处创建缓存区域；以及向所述缓存区域中写入由所述特效算法生成的多个算法参数。

根据本公开的一个示例实现方式，向缓存区域中写入多个算法参数包括：经由所述特效接口来从所述服务器设备获取所述多个算法参数中的目标算法参数；以及向缓

存区域中写入目标算法参数。

根据本公开的一个示例实现方式，获取目标算法参数包括：响应于确定在缓存区域中不存在目标算法参数，从服务器设备获取目标算法参数。

根据本公开的一个示例实现方式，生成目标图像帧包括：利用所述多个算法参数，
5 基于所述目标位置来生成所述目标图像帧。

根据本公开的一个示例实现方式，该方法进一步包括：响应于接收到用于退出编辑页面的退出请求，从客户端设备处移除特效接口。

根据本公开的一个示例实现方式，该方法进一步包括：响应于接收到用于切换输入图像的切换请求，清除缓存区域。

10 根据本公开的一个示例实现方式，从服务器处加载特效接口进一步包括：响应于确定特效算法为远程类型，从服务器设备加载特效接口。

根据本公开的一个示例实现方式，该方法进一步包括：响应于确定特效接口的加载状态为失败，呈现指示加载失败的提示信息。

15 示例装置和设备

图 8 示出了根据本公开的一些实现方式的用于生成视频的装置 800 的框图。该装置包括：第一呈现模块 810，被配置用于在客户端设备处呈现用于处理输入图像的处理页面；加载模块 820，被配置用于响应于接收到针对所述处理页面的加载请求，加载特效接口，所述特效接口用于调用在与所述客户端设备相对应的服务器设备处的特效
20 算法，来基于所述输入图像来生成具有预定特效的输出视频；以及第二呈现模块 830，被配置用于响应于确定特效接口的加载状态为成功，呈现用于编辑输出视频的编辑页面。

根据本公开的一个示例实现方式，编辑页面包括：用于呈现输出视频的预览区域、以及用于管理输出视频的播放的管理控件。

25 根据本公开的一个示例实现方式，该装置进一步包括：确定模块，被配置用于响应于接收到针对管理控件的管理请求，确定将被呈现在预览区域中的目标图像帧在输出视频中的目标位置；生成模块，被配置用于利用特效接口，生成对应于目标位置的目标图像帧；以及图像呈现模块，被配置用于在预览区域中呈现目标图像帧。

30 根据本公开的一个示例实现方式，该装置进一步包括：第三呈现模块，被配置用于呈现指示特效接口的加载进度的加载页面。

根据本公开的一个示例实现方式，加载模块包括：创建模块，被配置用于在客户端设备处创建缓存区域；以及写入模块，被配置用于向所述缓存区域中写入由所述特效算法生成的多个算法参数。

根据本公开的一个示例实现方式，写入模块包括：参数获取模块，被配置用于经由所述特效接口来从所述服务器设备获取所述多个算法参数中的目标算法参数；以及参数写入模块，被配置用于向缓存区域中写入目标算法参数。

5 根据本公开的一个示例实现方式，参数获取模块包括：远程获取模块，被配置用于响应于确定在缓存区域中不存在目标算法参数，从服务器设备获取目标算法参数。

根据本公开的一个示例实现方式，生成模块包括：读取模块，被配置用于利用所述多个算法参数，基于所述目标位置来生成所述目标图像帧。

根据本公开的一个示例实现方式，该装置进一步包括：移除模块，被配置用于响应于接收到用于退出编辑页面的退出请求，从客户端设备处移除特效接口。

10 根据本公开的一个示例实现方式，该装置进一步包括：切换模块，被配置用于响应于接收到用于切换输入图像的切换请求，清除缓存区域。

根据本公开的一个示例实现方式，加载模块包括：远程加载模块，被配置用于响应于确定特效算法为远程类型，从服务器设备加载特效接口。

15 根据本公开的一个示例实现方式，该装置进一步包括：第四呈现模块，被配置用于响应于确定特效接口的加载状态为失败，呈现指示加载失败的提示信息。

图9示出了能够实施本公开的多个实现方式的设备900的框图。应当理解，图9所示出的计算设备900仅仅是示例性的，而不应当构成对本文所描述的实现方式的功能和范围的任何限制。图9所示出的计算设备900可以用于实现上文描述的方法。

20 如图9所示，计算设备900是通用计算设备的形式。计算设备900的组件可以包括但不限于一个或多个处理器或处理单元910、存储器920、存储设备930、一个或多个通信单元940、一个或多个输入设备950以及一个或多个输出设备960。处理单元910可以是实际或虚拟处理器并且能够根据存储器920中存储的程序来执行各种处理。在多处理器系统中，多个处理单元并行执行计算机可执行指令，以提高计算设备900的并行处理能力。

25 计算设备900通常包括多个计算机存储介质。这样的介质可以是计算设备900可访问的任何可以获得的介质，包括但不限于易失性和非易失性介质、可拆卸和不可拆卸介质。存储器920可以是易失性存储器（例如寄存器、高速缓存、随机访问存储器（RAM））、非易失性存储器（例如，只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、闪存）或它们的某种组合。存储设备930可以是可拆卸或不可拆卸的介质，并且可以包括机器可读介质，诸如闪存驱动、磁盘或者任何其他介质，其
30 可以能够用于存储信息和/或数据（例如用于训练的训练数据）并且可以在计算设备900内被访问。

计算设备900可以进一步包括另外的可拆卸/不可拆卸、易失性/非易失性存储介

质。尽管未在图 9 中示出，可以提供用于从可拆卸、非易失性磁盘（例如“软盘”）进行读取或写入的磁盘驱动和用于从可拆卸、非易失性光盘进行读取或写入的光盘驱动。在这些情况中，每个驱动可以由一个或多个数据介质接口被连接至总线（未示出）。存储器 920 可以包括计算机程序产品 925，其具有一个或多个程序模块，这些程序模块被配置为执行本公开的各种实现方式的各种方法或动作。

通信单元 940 实现通过通信介质与其他计算设备进行通信。附加地，计算设备 900 的组件的功能可以以单个计算集群或多个计算机器来实现，这些计算机器能够通过通信连接进行通信。因此，计算设备 900 可以使用与一个或多个其他服务器、网络个人计算机（PC）或者另一个网络节点的逻辑连接来在联网环境中进行操作。

输入设备 950 可以是一个或多个输入设备，例如鼠标、键盘、追踪球等。输出设备 960 可以是一个或多个输出设备，例如显示器、扬声器、打印机等。计算设备 900 还可以根据需要通过通信单元 940 与一个或多个外部设备（未示出）进行通信，外部设备诸如存储设备、显示设备等，与一个或多个使得用户与计算设备 900 交互的设备进行通信，或者与使得计算设备 900 与一个或多个其他计算设备通信的任何设备（例如，网卡、调制解调器等）进行通信。这样的通信可以经由输入/输出（I/O）接口（未示出）来执行。

根据本公开的示例性实现方式，提供了一种计算机可读存储介质，其上存储有计算机可执行指令，其中计算机可执行指令被处理器执行以实现上文描述的方法。根据本公开的示例性实现方式，还提供了一种计算机程序产品，计算机程序产品被有形地存储在非瞬态计算机可读介质上并且包括计算机可执行指令，而计算机可执行指令被处理器执行以实现上文描述的方法。根据本公开的示例性实现方式，提供了一种计算机程序产品，其上存储有计算机程序，程序被处理器执行时实现上文描述的方法。

这里参照根据本公开实现的方法、装置、设备和计算机程序产品的流程图和/或框图描述了本公开的各个方面。应当理解，流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合，都可以由计算机可读程序指令实现。

这些计算机可读程序指令可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理单元，从而生产出一种机器，使得这些指令在通过计算机或其他可编程数据处理装置的处理单元执行时，产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中，这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作，从而，存储有指令的计算机可读介质则包括一个制造品，其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

可以把计算机可读程序指令加载到计算机、其他可编程数据处理装置、或其他设

备上，使得在计算机、其他可编程数据处理装置或其他设备上执行一系列操作步骤，以产生计算机实现的过程，从而使得在计算机、其他可编程数据处理装置、或其他设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

5 附图中的流程图和框图显示了根据本公开的多个实现的系统、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、程序段或指令的一部分，模块、程序段或指令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。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，两个连续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
10 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动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。

以上已经描述了本公开的各实现，上述说明是示例性的，并非穷尽性的，并且也不限于所公开的各实现。在不偏离所说明的各实现的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，
15 旨在最好地解释各实现的原理、实际应用或对市场中的技术的改进，或者使本技术领域的其他普通技术人员能理解本文公开的各个实现方式。

权利要求书

1. 一种用于生成视频的方法，包括：

在客户端设备处呈现用于处理输入图像的处理页面；

5 响应于接收到针对所述处理页面的加载请求，加载特效接口，所述特效接口用于调用在与所述客户端设备相对应的服务器设备处的特效算法，来基于所述输入图像来生成具有预定特效的输出视频；以及

响应于确定所述特效接口的加载状态为成功，呈现用于编辑所述输出视频的编辑页面。

10 2. 根据权利要求 1 所述的方法，其中所述编辑页面包括：用于呈现所述输出视频的预览区域、以及用于管理所述输出视频的播放的管理控件。

3. 根据权利要求 2 所述的方法，进一步包括：

响应于接收到针对所述管理控件的管理请求，确定将被呈现在所述预览区域中的目标图像帧在所述输出视频中的目标位置；

15 利用所述特效接口，生成对应于所述目标位置的所述目标图像帧；以及
在所述预览区域中呈现所述目标图像帧。

4. 根据权利要求 1 所述的方法，进一步包括：呈现指示所述特效接口的加载进度的加载页面。

5. 根据权利要求 3 所述的方法，其中加载所述特效接口包括：

20 在所述客户端设备处创建缓存区域；以及
向所述缓存区域中写入由所述特效算法生成的多个算法参数。

6. 根据权利要求 5 所述的方法，其中向所述缓存区域中写入所述多个算法参数包括：

25 经由所述特效接口来从所述服务器设备获取所述多个算法参数中的目标算法参数；
以及

向所述缓存区域中写入所述目标算法参数。

7. 根据权利要求 6 所述的方法，其中获取所述目标算法参数包括：响应于确定在所述缓存区域中不存在所述目标算法参数，从所述服务器设备获取所述目标算法参数。

30 8. 根据权利要求 3 所述的方法，其中生成所述目标图像帧包括：利用所述多个算法参数，基于所述目标位置来生成所述目标图像帧。

9. 根据权利要求 1 所述的方法，进一步包括：响应于接收到用于退出所述编辑页面的退出请求，从客户端设备处移除所述特效接口。

10. 根据权利要求 5 所述的方法，进一步包括：响应于接收到用于切换所述输入

图像的切换请求，清除所述缓存区域。

11. 根据权利要求 1 所述的方法，其中从服务器处加载所述特效接口进一步包括：
响应于确定所述特效算法为远程类型，从所述服务器设备加载所述特效接口。

5 12. 根据权利要求 1 所述的方法，进一步包括：响应于确定所述特效接口的所述
加载状态为失败，呈现指示加载失败的提示信息。

13. 一种用于生成视频的装置，包括：

第一呈现模块，被配置用于在客户端设备处呈现用于处理输入图像的处理页面；

10 加载模块，被配置用于响应于接收到针对所述处理页面的加载请求，加载特效接口，所述特效接口用于调用在与所述客户端设备相对应的服务器设备处的特效算法，
来基于所述输入图像来生成具有预定特效的输出视频；以及

第二呈现模块，被配置用于响应于确定所述特效接口的加载状态为成功，呈现用于编辑所述输出视频的编辑页面。

14. 一种电子设备，包括：

至少一个处理单元；以及

15 至少一个存储器，所述至少一个存储器被耦合到所述至少一个处理单元并且存储
用于由所述至少一个处理单元执行的指令，所述指令在由所述至少一个处理单元执行
时使所述电子设备执行根据权利要求 1 至 12 中任一项所述的方法。

15 15. 一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序在被处
理器执行时使所述处理器实现根据权利要求 1 至 12 中任一项所述的方法。

20

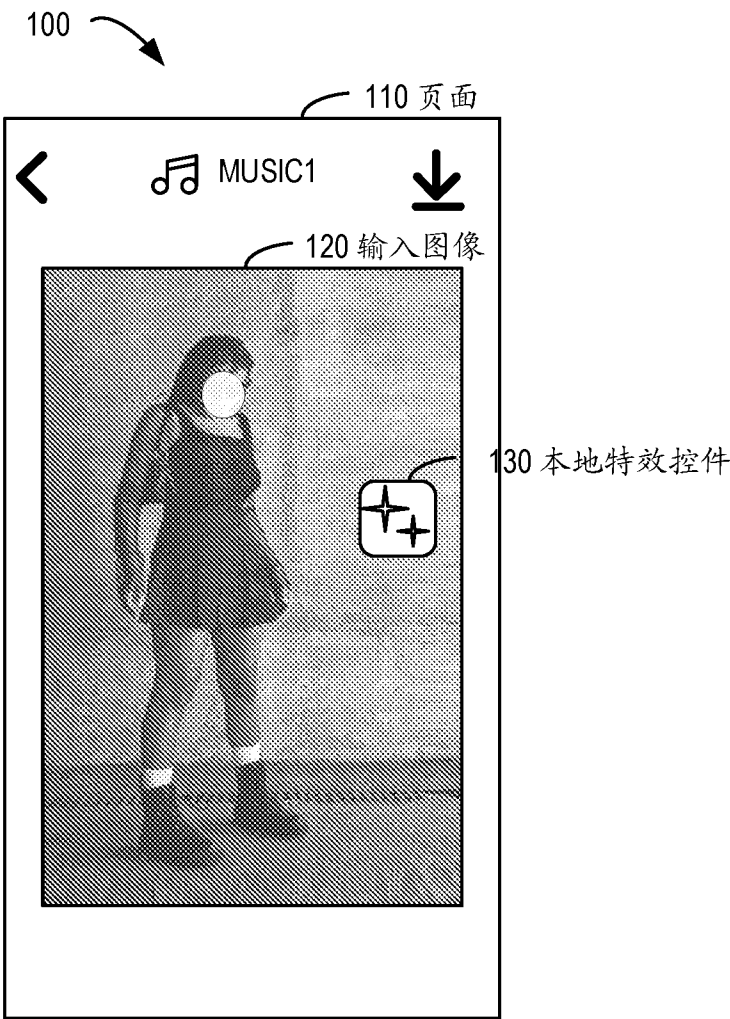


图1

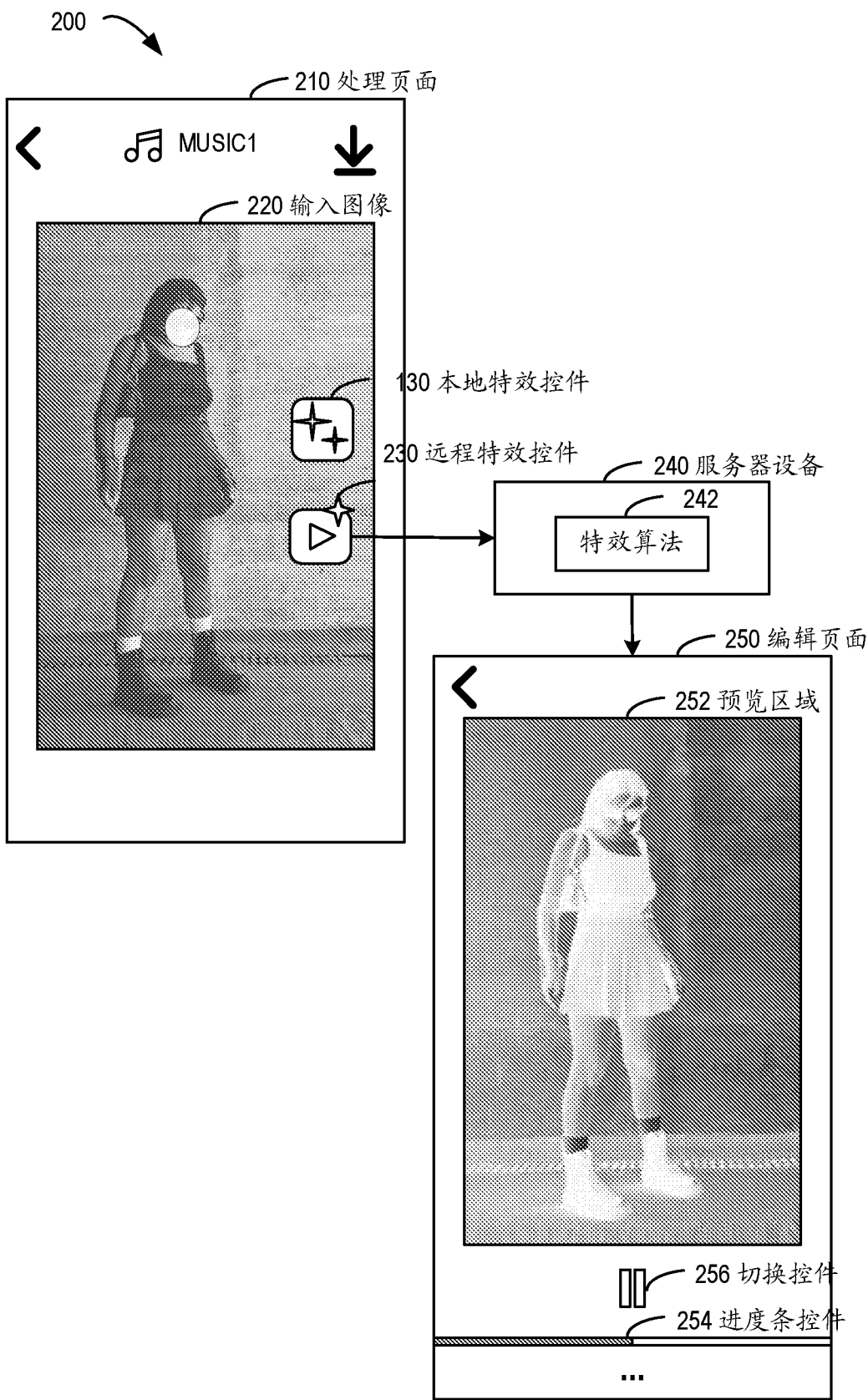


图2



图3



图4



图5

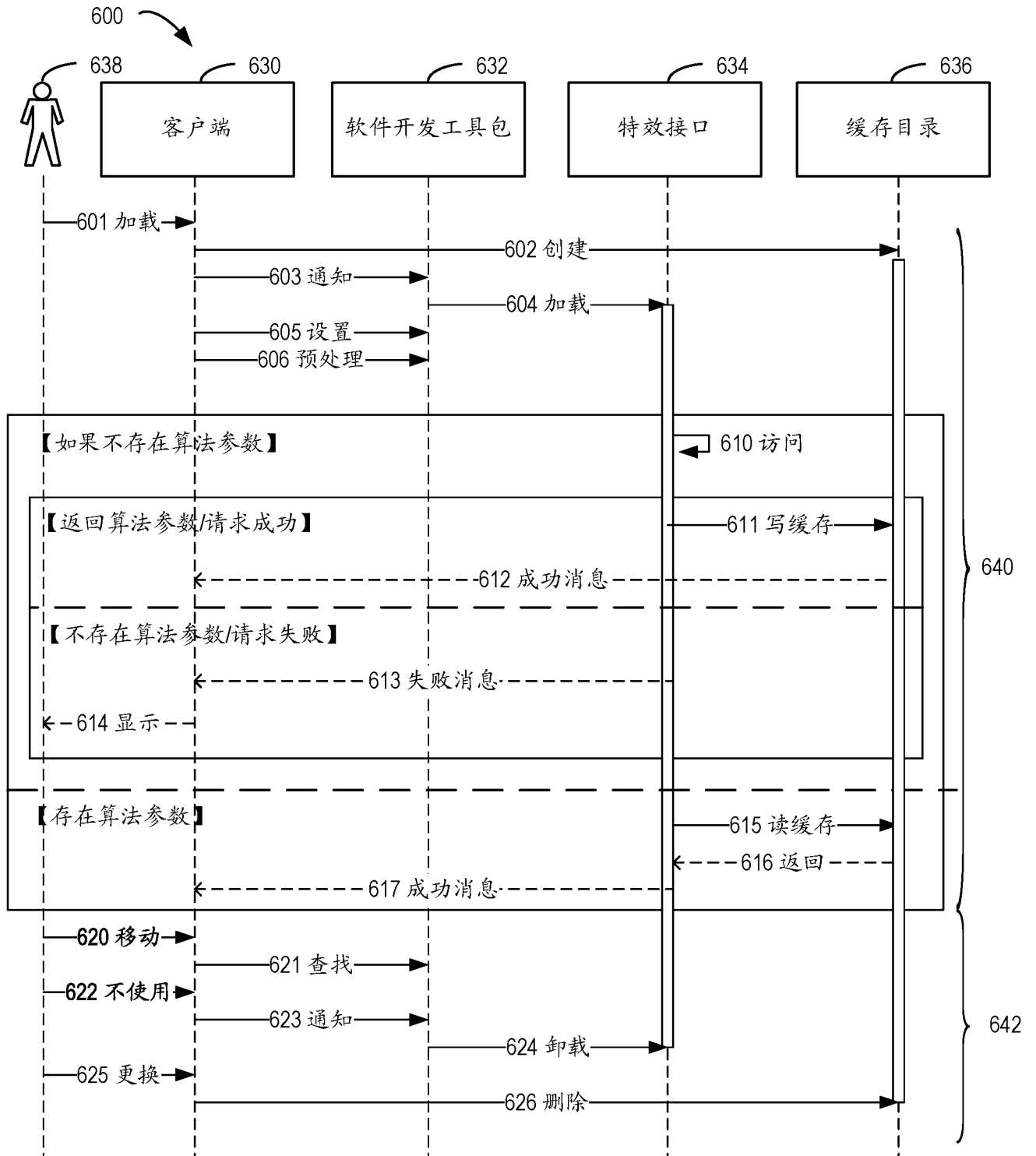


图6

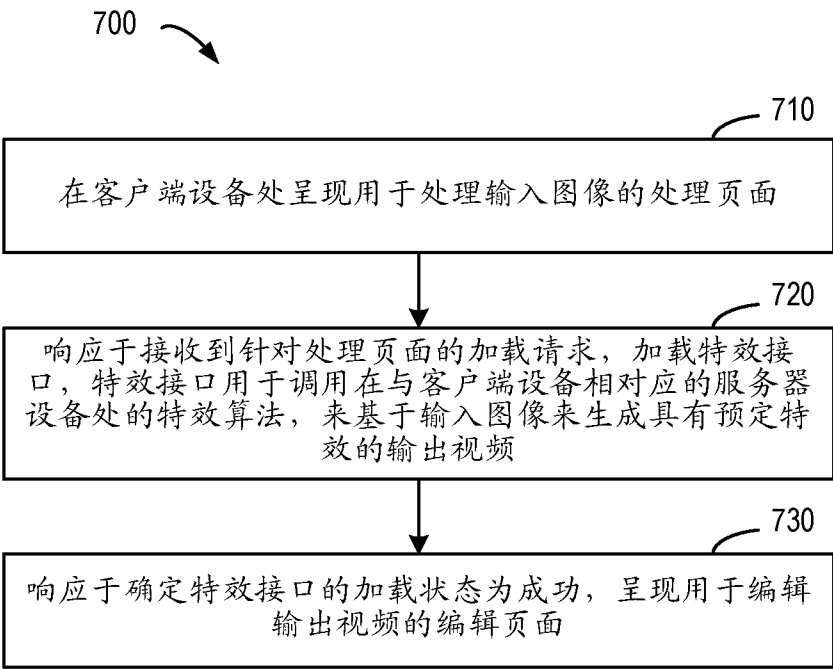


图7

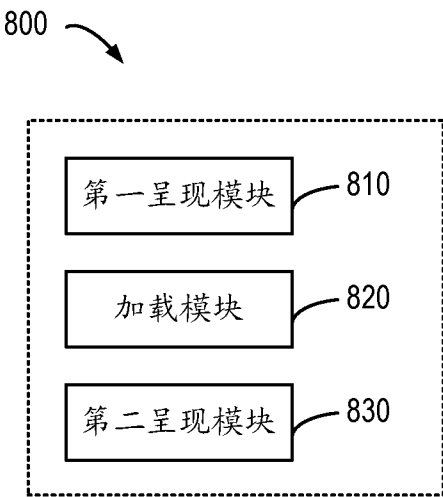


图8

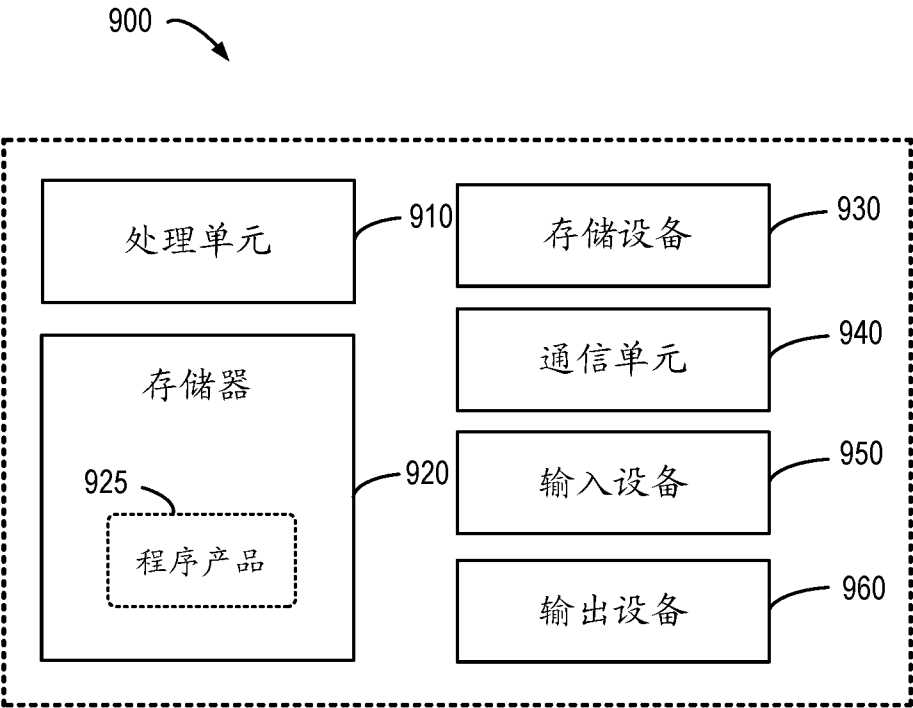


图9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/138191

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/472(2011.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC:H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CJFD, VEN, WOTXT, USTXT, EPTXT, IEEE: 视频, 特效, 算法, 参数, 服务器, 云, 添加, 加载, 缓存, 进度条; video, special effect, algorithm, parameter, server, cloud, add, cache, progress bar

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 114125551 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 01 March 2022 (2022-03-01) description, paragraphs [0069]-[0170], and figures 4-6	1-15
Y	CN 116934887 A (FACEMON CO., LTD.) 24 October 2023 (2023-10-24) description, paragraphs [0036]-[0154], and figures 1-14	1-15
Y	CN 108022279 A (GUANGZHOU BAIGUOYUAN INFORMATION TECHNOLOGY CO., LTD.) 11 May 2018 (2018-05-11) description, paragraphs [0120]-[0150], and figures 5-9	2, 3, 5-8, 10
PX	CN 117768735 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD. et al.) 26 March 2024 (2024-03-26) claims 1-15	1-15
A	CN 116471450 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 21 July 2023 (2023-07-21) entire document	1-15



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

20 March 2025

Date of mailing of the international search report

22 March 2025

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/138191

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2021196890 A1 (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 07 October 2021 (2021-10-07) entire document	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2024/138191

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	114125551	A	01 March 2022	None			
CN	116934887	A	24 October 2023	None			
CN	108022279	A	11 May 2018	WO	2019105438	A1	06 June 2019
CN	117768735	A	26 March 2024	None			
CN	116471450	A	21 July 2023	None			
WO	2021196890	A1	07 October 2021	None			

A. 主题的分类

H04N 21/472(2011.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC:H04N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, CNTXT, CJFD, VEN, WOTXT, USTXT, EPTXT, IEEE: 视频, 特效, 算法, 参数, 服务器, 云, 添加, 加载, 缓存, 进度条; video, special effect, algorithm, parameter, server, cloud, add, cache, progress bar

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 114125551 A (北京字节跳动网络技术有限公司) 2022年3月1日 (2022 - 03 - 01) 说明书第[0069]-[0170]段, 图4-6	1-15
Y	CN 116934887 A (脸萌有限公司) 2023年10月24日 (2023 - 10 - 24) 说明书第[0036]-[0154]段, 图1-14	1-15
Y	CN 108022279 A (广州市百果园信息技术有限公司) 2018年5月11日 (2018 - 05 - 11) 说明书第[0120]-[0150]段, 图5-9	2, 3, 5-8, 10
PX	CN 117768735 A (北京字跳网络技术有限公司等) 2024年3月26日 (2024 - 03 - 26) 权利要求1-15	1-15
A	CN 116471450 A (腾讯科技(深圳)有限公司) 2023年7月21日 (2023 - 07 - 21) 全文	1-15
A	WO 2021196890 A1 (TENCENT TECH SHENZHEN CO LTD) 2021年10月7日 (2021 - 10 - 07) 全文	1-15

☐ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2025年3月20日

国际检索报告邮寄日期

2025年3月22日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

陈德锋

电话号码 (+86) 010-62411501

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/138191

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	114125551	A	2022年3月1日	无	
CN	116934887	A	2023年10月24日	无	
CN	108022279	A	2018年5月11日	WO 2019105438 A1	2019年6月6日
CN	117768735	A	2024年3月26日	无	
CN	116471450	A	2023年7月21日	无	
WO	2021196890	A1	2021年10月7日	无	