

(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
2022 年 12 月 8 日 (08.12.2022)



(10) 国际公布号
WO 2022/253276 A1

(51) 国际专利分类号:
H04N 21/44 (2011.01)

(21) 国际申请号: PCT/CN2022/096646

(22) 国际申请日: 2022 年 6 月 1 日 (01.06.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110610952.5 2021年6月1日 (01.06.2021) 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, Beijing 100190 (CN)。

(72) 发明人: 何沃洲 (**HE, Wozhou**); 中国北京市海淀区知春路63号中国卫星通信大厦今日
头条小邮局, Beijing 100086 (CN)。 李根 (**LI,**

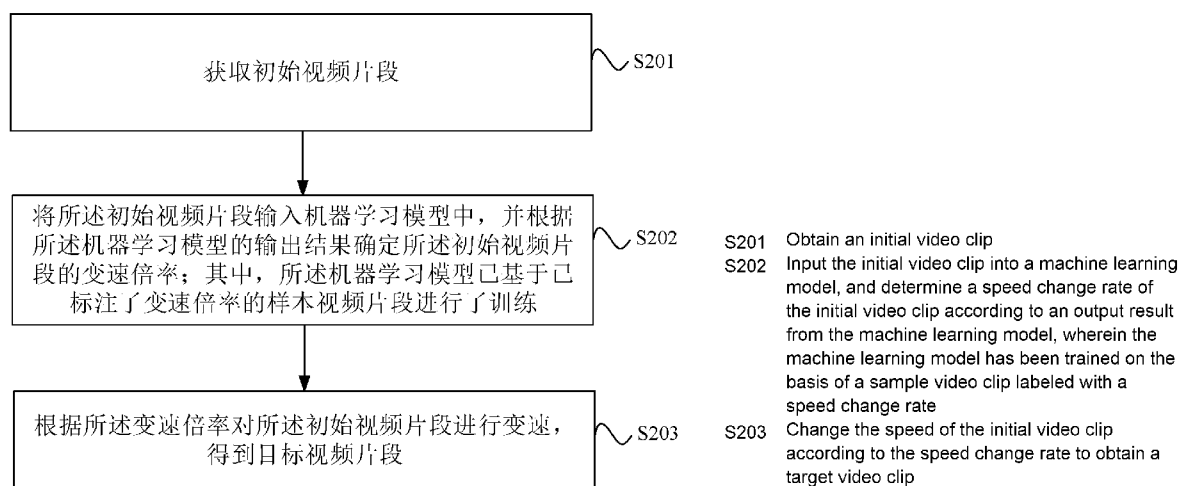
Gen); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100086 (CN)。 吴兴龙 (**WU, Xinglong**); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100086 (CN)。

(74) 代理人: 上海光栅知识产权代理有限公司
(**LINKER IP LLC**); 中国上海市长宁区延安西路
889号702室, Shanghai 201109 (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,
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, ME, MG, MK, MN,
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,

(54) **Title:** VIDEO PROCESSING METHOD AND DEVICE, STORAGE MEDIUM, AND COMPUTER PROGRAM PRODUCT

(54) 发明名称: 视频处理方法、设备、存储介质及计算机程序产品



(57) **Abstract:** Embodiments of the present disclosure provide a video processing method and device, a storage medium, and a computer program product. The method comprises: obtaining an initial video clip (S201); inputting the initial video clip into a machine learning model, and determining a speed change rate of the initial video clip according to an output result from the machine learning model, wherein the machine learning model has been trained on the basis of a sample video clip labeled with a speed change rate (S202); and changing the speed of the initial video clip according to the speed change rate to obtain a target video clip (S203). According to the embodiments of the present disclosure, by determining a speed change rate of an initial video clip by a machine learning model by using artificial intelligence, the reliance on editing experience of a user is reduced; user operations are simplified; processing efficiency is improved; labor cost is reduced; the performance capability of a video is effectively improved; information density of a video is reasonably improved; and the overall number of playbacks of a video after being published is increased.

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, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, 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, 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))。

(57) 摘要: 本公开实施例提供一种视频处理方法、设备、存储介质及计算机程序产品, 通过获取初始视频片段(S201); 将初始视频片段输入机器学习模型中, 并根据机器学习模型的输出结果确定初始视频片段的变速倍率; 其中, 机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练(S202); 根据变速倍率对初始视频片段进行变速, 得到目标视频片段(S203)。本公开实施例应用人工智能的方式通过机器学习模型来确定初始视频片段的变速倍率, 降低对用户编辑经验的依赖, 简化用户操作, 提高处理效率, 降低人力成本, 且有效地提高视频表现能力, 合理地提高视频信息密度, 提高视频发布后的整体播放数。

视频处理方法、设备、存储介质及计算机程序产品

相关申请交叉引用

5 本申请要求于 2021 年 6 月 1 日提交中国专利局、申请号为 202110610952.5、发明名称为“视频处理方法、设备、存储介质及计算机程序产品”的中国专利申请的优先权，其全部内容通过引用并入本文。

技术领域

10 本公开实施例涉及计算机与网络通信技术领域，尤其涉及一种视频处理方法、设备、存储介质及计算机程序产品。

背景技术

15 视频变速是一种常用的视频编辑手段，在进行视频编辑时，用户可以选择视频素材的某个片段，人工对视频片段进行某一变速倍率的常规变速，或者制定曲线形状的曲线变速。然而，通过人工方式对视频进行变速编辑往往需要用户具有一定的经验，经验不足的用户可能尝试若干次才能确定相对合适的变速效果，并且用户确定变速倍率的效率较低，最终确定的变速倍率可能达不到最佳的变速效果，使得最终得到的视频表现能力不佳。

发明内容

20 本公开实施例提供一种视频处理方法、设备、存储介质及计算机程序产品，以准确确定视频片段的变速倍率，实现视频智能变速，简化用户操作，不需要用户具有编辑经验，提高视频表现能力。

第一方面，本公开实施例提供一种视频处理方法，包括：

获取初始视频片段；

25 将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率；其中，所述机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；

根据所述变速倍率对所述初始视频片段进行变速，得到目标视频片段。

第二方面，本公开实施例提供一种模型训练方法，包括：

30 获取多个训练数据；所述训练数据包括已标注了变速倍率的样本视频片段；

根据所述训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

第三方面，本公开实施例提供一种视频处理设备，包括：

获取单元，用于获取初始视频片段；

35 处理单元，用于将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率；其中，所述机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；

变速单元，用于根据所述变速倍率对所述初始视频片段进行变速，得到目标视频片段。

第四方面，本公开实施例提供一种模型训练设备，包括：

获取单元，用于获取多个训练数据；所述训练数据包括已标注了变速倍率的样本视频片段；

训练单元，用于根据所述训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

- 5 第五方面，本公开实施例提供一种电子设备，包括：至少一个处理器和存储器；
所述存储器存储计算机执行指令；

所述至少一个处理器执行所述存储器存储的计算机执行指令，使得所述至少一个处理器执行如上第一方面、第二方面以及第一方面、第二方面各种可能的设计所述的方法。

- 10 第六方面，本公开实施例提供一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机执行指令，当处理器执行所述计算机执行指令时，实现如上第一方面、第二方面以及第一方面、第二方面各种可能的设计所述的方法。

第七方面，本公开实施例提供一种计算机程序产品，包括计算机指令，该计算机指令被处理器执行时实现如上第一方面、第二方面以及第一方面、第二方面各种可能的设计所述的方法。

- 15 第八方面，本公开实施例提供一种计算机程序，该计算机程序被处理器执行时，实现如上第一方面、第二方面以及第一方面、第二方面各种可能的设计所述的方法。

- 本实施例提供的视频处理方法、设备、存储介质及计算机程序产品，通过获取初始视频片段；将初始视频片段输入机器学习模型中，并根据机器学习模型的输出结果确定初始视频片段的变速倍率；其中，机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；
20 根据变速倍率对初始视频片段进行变速，得到目标视频片段。本公开实施例应用人工智能的方式通过机器学习模型来确定初始视频片段的变速倍率，降低对用户编辑经验的依赖，简化用户操作，提高处理效率，降低人力成本，且有效地提高视频表现能力，合理地提高视频信息密度，提高视频发布后的整体播放数。

25 附图说明

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一简单地介绍，显而易见地，下面描述中的附图是本公开的一些实施例，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。

- 30 图1为本公开实施例提供的视频处理方法的应用场景示例图；
图2为本公开一实施例提供的视频处理方法流程示意图；
图3为本公开另一实施例提供的视频处理方法流程示意图；
图4为本公开另一实施例提供的视频处理方法流程示意图；
图5为本公开另一实施例提供的视频处理方法流程示意图；
35 图6为本公开一实施例提供的视频处理设备的结构框图；
图7为本公开一实施例提供的模型训练设备的结构框图；
图8为本公开一实施例提供的电子设备的硬件结构示意图。

具体实施方式

为使本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本公开一部分实施例，而不是全部的实施例。基于本公开中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本公开保护的范围。

5 现有技术中，在进行视频编辑时，通常需要用户选择视频素材的某个片段，人工对视频片段进行某一变速倍率的常规变速，或者制定曲线形状的曲线变速，也即变速倍率呈现曲线分布，其中通常情况下，变速倍率大于 1.0 为快放，小于 1.0 为慢放。

然而，现有技术中通过人工方式对视频进行变速编辑往往需要一定的经验，用户可能尝试若干次才能确定相对合适的变速倍率，达到相对合适的变速效果，并且用户确定变速倍率的效率较低，在面对较长的视频素材需要人工一段一段的设定变速倍率，消耗大量的人力成本和时间成本，且最终确定的变速倍率也可能达不到最佳的变速效果，使得最终得到的视频表现能力不佳，尤其是无法准确科学的提高视频的信息密度、提高视频发布后的整体播放数（Video View，简称 VV，指在一个统计周期内视频被打开的总次数）。

为了解决上述技术问题，本公开实施例中考虑应用人工智能的方式来确定视频素材的变速倍率，以准确确定初始视频各视频片段的变速倍率，降低对用户编辑经验的依赖，简化用户操作，提高处理效率，降低人力成本，且有效地提高视频表现能力，合理地提高视频信息密度，提高视频发布后的整体播放数。

具体的，可训练一个机器学习模型，该模型的输入为初始视频片段，根据机器学习模型的输出结果可确定初始视频片段的变速倍率，进而可根据变速倍率对初始视频片段进行变速。

20 本公开实施例的视频处理方法可应用于视频编辑过程中，处理后得到经过变速的目标视频。当然，该视频处理方法也可应用在视频播放过程中，也即在播放某一视频时，通过上述过程确定当前播放的视频的某一初始视频片段的变速倍率，按照该变速倍率对该初始视频片段进行播放。

本公开实施例的视频处理方法适用于如图 1 所示的应用场景，包括：终端设备 10，还可包括视频采集设备 12 和/或视频存储设备 13，其中视频采集设备 12 和/或视频存储设备 13 用于向终端设备 10 提供初始视频，或者若终端设备 10 配置有摄像头，则应用场景中也可仅包括终端设备 10，由终端设备 10 自己采集初始视频；终端设备 10 可包括但不限于智能手机、平板电脑、笔记本电脑、个人电脑、个人数字助理、可穿戴设备等等，终端设备 10 可将初始视频中得到初始视频片段，例如对初始视频进行分割，进而基于机器学习模型进行处理，根据机器学习模型的输出结果确定初始视频片段的变速倍率，根据变速倍率对初始视频片段进行变速，得到目标视频片段。

其中机器学习模型可预先配置在终端设备 10 上；当然也可配置在云端服务器上，终端设备 10 可将任一视频片段的视频帧发送给云端服务器，由云端服务器获取视频片段对应的评价参数后发送给终端设备 10。此外对于机器学习模型的训练过程可以预先在终端设备 10 上进行，当然也可预先在云端服务器上，再将训练好的模型配置在终端设备 10 上。

下面将结合具体实施例对本公开提供的视频处理方法流程进行详细介绍。

参考图 2，图 2 为本公开实施例提供的视频处理方法流程示意图。本实施例的方法可以应用在终端设备或服务器中，该视频处理方法包括：

S201、获取初始视频片段。

在本实施例中，初始视频片段可以为从视频采集设备或视频存储设备获取到的视频片段，也可以是本地存储的视频片段，视频片段可以是一整段视频，也可以是一整段视频中的某一个片段，此处可不做限制。

S202、将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率；其中，所述机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练。

在本实施例中，机器学习模型可以为神经网络模型，或者其他机器学习模型，此处可不做限制。此外，机器学习模型的输入为初始视频片段的视频文件，也可以为从初始视频片段中提取的视频帧；机器学习模型的输出可以为初始视频片段的变速倍率，也可以为初始视频片段的评价参数，评价参数为用于衡量初始视频片段适宜变速程度的大小的参数，基于初始视频片段的评价参数可确定初始视频片段的变速倍率，当然机器学习模型的输出也可不限于上述举例，可基于机器学习模型的输出得到初始视频片段的变速倍率即可。相应的，可根据输入和输出的不同设计和训练机器学习模型，此处不再赘述。

S203、根据所述变速倍率对所述初始视频片段进行变速，得到目标视频片段。

在本实施例中，在得到初始视频片段的变速倍率后，可根据变速倍率对初始视频片段进行变速处理。

可选的，在视频编辑过程中，可基于初始视频片段的变速倍率，完成对初始视频的编辑，从而生成新视频。具体的，初始视频片段来源于初始视频时，可在得到目标视频片段之后，通过在初始视频中将初始视频片段替换为所述目标视频片段，得到目标视频。

当然，上述过程也可应用在视频播放过程中，也即在播放初始视频片段时，根据初始视频片段的变速倍率对初始视频片段进行播放。

本实施例提供的视频处理方法，通过获取初始视频片段；将初始视频片段输入机器学习模型中，并根据机器学习模型的输出结果确定初始视频片段的变速倍率；其中，机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；根据变速倍率对初始视频片段进行变速，得到目标视频片段。本实施例应用人工智能的方式通过机器学习模型来确定初始视频片段的变速倍率，降低对用户编辑经验的依赖，简化用户操作，提高处理效率，降低人力成本，且有效地提高视频表现能力，合理地提高视频信息密度，提高视频发布后的整体播放数。

在上述实施例的基础上，S201所述的获取初始视频片段，包括：

S2021、获取初始视频；

S2022、通过对所述初始视频进行分镜，得到所述初始视频片段，其中，所述初始视频片段为所述初始视频中的分镜片段。

在本实施例中，对于初始视频，可能某些片段适宜变速，某些片段不适宜变速，因此可对初始视频进行分割，将初始视频分割成更小粒度的视频片段，以便于分别对每一段视频片段判断是否适宜变速，并确定适宜变速的视频片段的变速倍率。

可选的，本实施例考虑到对视频某一部分进行变速时该部分视频通常具有关联的部分，关联的部分通常属于同一个场景，因此在将初始视频分割为多个视频片段时，可根据场景对所述初始视频进行分镜，同一场景作为一个分镜片段，上述实施例中的初始视频片段可以为某一分镜片段。

可选的，考虑到一个分镜片段也可能较长，因此在分镜片段的基础上按照预设时长进行分割，得到多个视频片段，例如预设时长可以为 1 秒或者其他时长，本实施例中还可将分镜片段的任一视频片段作为初始视频片段。

5 可选的，在对初始视频的分镜片段整段进行编辑的过程中，需要获取分镜片段的每一视频片段的变速倍率。可选的，在对分镜片段分割时，相邻的视频片段可具有一定的时间交叠，以预设时长为 1 秒为例，本实施例中相邻的视频片段可具有一定的时间交叠，例如第一段视频片段可以为 0~1 秒，第二段视频片段可以为 0.5~1.5 秒，第三段视频片段可以为 1~2 秒，以此类推，通过时间交叠，可以在后续过程中更加准确地确定各视频片段适宜变速的程度，但处理量相对增加，仅此时间交叠的大小可根据处理量和效果等多种因素进行设定。当然，相
10 邻的视频片段也可不具有时间交叠，也即第一段视频片段可以为 0~1 秒，第二段视频片段可以为 1~2 秒，以此类推。此外，若分镜片段的时长小于或等于预设时长，则可不对分镜片段进行分割。

进一步的，在上述实施例的基础上，还可对初始视频片段进行抽帧，提取初始视频片段中的视频帧作为机器学习模型的输入，可选的，可以按照预设帧率对初始视频片段进行抽帧，
15 得到初始视频片段中的视频帧，例如预设帧率可以为 30fps，也即每一秒视频片段中可以抽取到 30 个视频帧。

在上述实施例的基础上，如图 3 所示，S202 所述的将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率，具体可包括：

S301、提取所述初始视频片段的视频帧；
20 S302、将所述初始视频片段的视频帧输入机器学习模型中，获取所述初始视频片段对应的评价参数，所述评价参数为用于衡量所述初始视频片段适宜变速程度的大小的参数；
S303、根据所述初始视频片段对应的评价参数、以及预设评价参数与变速倍率对照关系，确定所述初始视频片段的变速倍率。

在本实施例中，可基于机器学习模型来实现根据所述初始视频片段的视频帧，得到初始
25 视频片段的变速倍率。

可选的，该机器学习模型可以为评价初始视频片段适宜变速程度的神经网络模型；该模型的输入为初始视频片段的视频帧，输出为对该视频片段的评价参数，该评价参数用于表示初始视频片段适宜变速的程度，例如评价参数的取值范围为 0~1，越接近 0 表示对该初始视频片段越适宜慢放，越接近 1 表示对该初始视频片段越适宜快放。其中，适宜慢放的视频片段
30 通常具有如下至少一种特征：主体清晰、镜头或主体有较快速的运动、能展示出更多的细节和美感（例如倒水、快跑和击球的瞬间等等）；而适宜快放的视频片段通常相邻帧之间变化比较小、包含信息量比较少等等。通过训练神经网络模型进行学习，可以根据输入的初始视频片段的视频帧得到对应的适宜变速程度的评价参数。

进一步的，在获取到初始视频片段的评价参数后，可根据该初始视频片段的评价参数获
35 取初始视频片段的变速倍率。在上述实施例中，评价参数不同代表视频片段适宜变速的程度不同，尤其是，评价参数的取值范围为 0~1，越接近 0 表示对该视频片段越适宜慢放，也即评价参数越小对该视频片段越适宜慢放，越接近 1 表示对该视频片段越适宜快放，也即评价参数越大对该视频片段越适宜快放，评价参数与变速倍率存在对照关系，本实施例中可预先获取评价参数与变速倍率对照关系，进而可根据预设的评价参数与变速倍率对照关系、以及初

始视频片段对应的评价参数，获取初始视频片段的变速倍率，也即通过查评价参数与变速倍率对照关系，找到初始视频片段对应的评价参数对应的变速倍率，作为所述初始视频片段的变速倍率。

在另一种可选实施例中，该机器学习模型为获取变速倍率的神经网络模型，也即该模型的输入为初始视频片段的视频帧，输出为初始视频片段的变速倍率，也即不再单独执行评价参数到变速倍率之间的转换过程。因此基于该机器学习模型，S202 具体为将所述初始视频片段的视频帧输入机器学习模型，通过所述机器学习模型输出所述初始视频片段的变速倍率。

在上述任一实施例的基础上，可选的，在 S302 所述获取所述初始视频片段对应的评价参数后，还可包括：

对所 述初始视频片段对应的评价参数以及其相邻的其他初始视频片段对应的评价参数进行平滑处理，将平滑处理后的各评价参数确定为对应的各初始视频片段最终的评价参数。

在本实施例中，为了避免初始视频片段及其相邻视频片段在进行变速时存在突然变速、导致过渡不平滑不自然，影响视频表现能力，可以对相邻的视频片段的评价参数进行平滑处理，例如在时间轴维度上对评价参数进行加窗平滑处理，其中可以加三角窗或者矩形窗，使得评价参数随时间轴变化和过渡更加平滑和自然，进而可以使得后续获取到的初始视频片段及其相邻视频片段的变速倍率随时间轴变化和过渡也保持平滑和自然。

在另一种可选实施例中，可以在 S303 所述的确定所述初始视频片段的变速倍率后，对所述初始视频片段的变速倍率以及其相邻的其他初始视频片段的变速倍率进行平滑处理，将平滑处理后的各变速倍率确定为对应的各初始视频片段最终的变速倍率。通过本实施例对变速倍率进行平滑处理，也可以使得初始视频片段及其相邻视频片段的变速倍率随时间轴变化和过渡也保持平滑和自然。其中，平滑处理过程可同上述实施例，此处不再赘述。

在上述任一实施例的基础上，如图 4 所示，在确定所述初始视频片段的变速倍率后，还可包括：

S401、获取所述初始视频片段中的初始语速；

S402、根据所述初始语速以及所述初始视频片段的变速倍率，确定所述初始视频片段按照所述变速倍率播放时的目标语速；

S403、若所述目标语速超过预设语速阈值，则根据所述预设语速阈值调整所述变速倍率。

在本实施例中，由于一些初始视频片段可能伴随有语音，为了避免初始视频片段变速后语速过快导致信息量过分密集，或者语速过慢导致信息量过分稀疏，因此可以配置预设语速阈值，例如可以设置为平均语速 8 字/秒、峰值语速 10 字/秒等，进而在获取到初始视频片段中的初始语速后，根据初始语速以及该初始视频片段的目标变速语速，确定初始视频片段按照变速倍率播放时的目标语速，也即变速之后的语速，然后将目标语速与预设语速阈值进行比较，若目标语速超过预设语速阈值，则根据预设语速阈值调整变速倍率，使得初始视频片段按照调整后的变速倍率播放时的目标语速不超过预设语速阈值，并以调整后的变速倍率确定为初始视频片段的最终的变速倍率。

在上述任一实施例的基础上，如图 5 所示，本方法还包括对机器学习模型进行训练的过程，其执行主体可以为上述视频处理方法实施例的执行主体，也可以为其他任一终端设备或服务器等电子设备，其中机器学习模型为评价视频片段适宜变速程度的神经网络模型，训练过程具体如下：

S501、获取多个训练数据；所述训练数据包括已标注了变速倍率的样本视频片段；

S502、根据所述训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

在本实施例中，可以首先采集一些训练数据，所述训练数据可以包括已标注了变速倍率的样本视频片段，进而根据训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

可选的，可获取已标注了变速倍率的样本视频片段，所述样本视频片段为已被慢放或快放、且被标注有慢放标签或快放标签的视频素材，和/或，未被慢放或快放、且其各视频片段分别被标注有标识适宜变速程度的标签的视频素材；对所述样本视频片段提取视频帧，将所述视频帧确定为所述训练数据。其中，对于已被慢放或快放、且被标注有慢放标签或快放标签的视频素材，可以直接标注真实的变速倍率，而对于未被慢放或快放、且其各视频片段分别被标注有标识适宜变速程度的标签的视频素材，可以根据适宜变速程度赋予其合适的变速倍率。

可选的，训练数据还可以为标注预设标签的连续视频帧序列，其中，可选的，所述预设标签可包括标识适宜变速程度的标签或者标识变速情况的标签，标识变速情况的标签可以为标识已被慢放的标签、标识已被快放的标签，进一步还可标注对应的变速倍率；而标识适宜变速程度的标签可以是对适宜慢放的连续视频帧序列按照适宜程度赋予数值，对适宜快放的连续视频帧序列按照适宜程度赋予数值，作为视频片段的标签，作为示例，可以以 0、1、2、3 四个数值作为标签，其中，数值越大表示越适宜慢放，数值越小表示越适宜快放。

在一种可选实施例中，连续视频帧序列是通过从视频素材中提取的，具体的，可获取经标注的视频素材，经过标注的视频素材可以包括：已被慢放或快放、且被标注有慢放标签或快放标签的视频素材，和/或，未被慢放或快放、且其各视频片段分别被标注有适宜变速程度标签的视频素材，进一步的，可对所述视频素材的视频片段提取视频帧，将所述视频帧确定为所述训练数据。其中提取的视频帧可以携带视频素材标注的标签。通过已被慢放或快放、且被标注有慢放标签或快放标签的视频素材得到的视频帧进行模型训练，可以使得模型能够更好的学习到真实的被慢放或快放的视频片段具有怎样的特征；而通过未被慢放或快放、且其各视频片段分别被标注有适宜变速程度标签的视频素材得到的视频帧进行模型训练，可以使得模型能够更好的学习到什么样的视频片段适宜变速、以及对应的适宜变速程度。

其中视频素材标注过程可以人为进行，根据预定的标注规则进行标注，例如对于视频素材中主体清晰、镜头或主体有较快速的运动、能展示出更多的细节和美感（例如倒水、快跑和击球的瞬间等等）的视频片段，可以确定为适宜慢放的视频片段，而相邻帧之间变化比较小、包含信息量比较少的视频片段，可以确定为适宜快放的视频片段，可对适宜慢放的视频片段按照适宜程度赋予数值，对适宜快放的视频片段按照适宜程度赋予数值，作为视频片段的标签，作为示例，可以以 0、1、2、3 四个数值作为标签，其中，数值越大表示越适宜慢放，数值越小表示越适宜快放。

在获取到训练数据后，可对机器学习模型的初始模型进行训练，得到所述机器学习模型。其中机器学习模型的初始模型可以为 3D 卷积神经网络，或者也可以为其他人工智能模型、机器学习模型，此处不做限定，具体的训练过程此处也不做赘述。需要说明的是，本实施例中对于经过标注的视频素材可以进行视频片段的分割，例如先进行分镜，并将分镜片段按照预

设时长进行分割，从而得到多个视频片段，进而对每一视频片段进行抽帧，得到每一视频片段的视频帧。

本实施例中在训练机器学习模型的初始模型时，不同视频素材的视频片段的连续视频帧序列之间不限顺序，可以随机打乱，例如可以先根据第一视频素材的第二视频片段的连续视频帧序列训练模型，再根据第二视频素材的第一视频片段的连续视频帧序列训练模型，再根据第一视频素材的第一视频片段的连续视频帧序列训练模型等等，不限制训练顺序；当然本实施例中也可有序回归的训练，也即先根据第一视频素材的第一视频片段的连续视频帧序列训练模型，再根据第一视频素材的第二视频片段的连续视频帧序列训练模型，然后再根据第一视频素材的第三视频片段的连续视频帧序列训练模型，以此类推按顺序训练。在上述示例中，由于视频片段的标签为离散的数值，通过有序回归的结果时得到大于某数值的概率，例如对于某一视频片段，模型可以输出其大于 1 的概率是 0.8，大于 2 的概率是 0.5，大于 3 的概率是 0.2，然后将这些概率通过加权变换后得到 0~1 范围的评价参数。

通过上述各实施例提供的视频处理方法，应用人工智能的方式通过机器学习模型来确定初始视频的各视频片段的变速倍率，降低对用户编辑经验的依赖，简化用户操作，提高效率，降低人力成本，且有效地提高视频表现能力，合理地提高视频信息密度，提高视频发布后的整体播放数。

对应于上文实施例的视频处理方法，图 6 为本公开实施例提供的视频处理设备的结构框图。为了便于说明，仅示出了与本公开实施例相关的部分。参照图 6，所述视频处理设备 600 包括：获取单元 601、处理单元 602、变速单元 603。

获取单元 601，用于获取初始视频片段；

处理单元 602，用于将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率；其中，所述机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；

变速单元 603，用于根据所述变速倍率对所述初始视频片段进行变速，得到目标视频片段。

根据本公开的一个或多个实施例，所述获取单元 601 在获取初始视频片段时，用于：获取初始视频；

通过对所述初始视频进行分镜，得到所述初始视频片段，其中，所述初始视频片段为所述初始视频中的分镜片段。

根据本公开的一个或多个实施例，所述变速单元 603 在得到目标视频片段之后，还用于：通过在所述初始视频中将所述初始视频片段替换为所述目标视频片段，得到目标视频。

根据本公开的一个或多个实施例，所述获取单元 601 在通过对所述初始视频进行分镜，得到所述初始视频片段时，用于：

对所述初始视频进行分镜，得到所述初始视频中的分镜片段；

将所述分镜片段分割为多个视频片段，将其中任一视频片段确定为所述初始视频片段。

根据本公开的一个或多个实施例，所述处理单元 602 在将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率时，用于：提取所述初始视频片段的视频帧；

将所述初始视频片段的视频帧输入机器学习模型中，获取所述初始视频片段对应的评价参数，所述评价参数为用于衡量所述初始视频片段适宜变速程度的大小的参数；

根据所述初始视频片段对应的评价参数、以及预设评价参数与变速倍率对照关系，确定所述初始视频片段的变速倍率。

5 根据本公开的一个或多个实施例，所述处理单元 602 在获取所述初始视频片段对应的评价参数后，还用于：

对所述初始视频片段对应的评价参数以及其相邻的其他初始视频片段对应的评价参数进行平滑处理，将平滑处理后的各评价参数确定为对应的各初始视频片段最终的评价参数；或者

10 在所述确定所述初始视频片段的变速倍率后，还包括：

对所述初始视频片段的变速倍率以及其相邻的其他初始视频片段的变速倍率进行平滑处理，将平滑处理后的各变速倍率确定为对应的各初始视频片段最终的变速倍率。

根据本公开的一个或多个实施例，所述处理单元 602 在确定所述初始视频片段的变速倍率后，还用于：

15 获取所述初始视频片段中的初始语速；

根据所述初始语速以及所述初始视频片段的变速倍率，确定所述初始视频片段按照所述变速倍率播放时的目标语速；

若所述目标语速超过预设语速阈值，则根据所述预设语速阈值调整所述变速倍率。

20 本实施例提供的视频处理设备，可用于执行上述视频处理方法实施例的技术方案，其实现原理和技术效果类似，本实施例此处不再赘述。

对应于上文实施例的视频处理方法，图 7 为本公开实施例提供的视频处理设备的结构框图。为了便于说明，仅示出了与本公开实施例相关的部分。参照图 7，模型训练设备 610 所述设备包括：获取单元 611 以及训练单元 612。

25 获取单元 611，用于获取多个训练数据；所述训练数据包括已标注了变速倍率的样本视频片段；

训练单元 612，用于根据所述训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

根据本公开的一个或多个实施例，所述获取单元 611 在获取多个训练数据时，用于：

30 获取已标注了变速倍率的样本视频片段，所述样本视频片段为已被慢放或快放、且被标注有慢放标签或快放标签的视频素材，和/或，未被慢放或快放、且其各视频片段分别被标注有标识适宜变速程度的标签的视频素材；

对所述样本视频片段提取视频帧，将所述视频帧确定为所述训练数据。本实施例提供的模型训练设备，可用于执行上述模型训练方法实施例的技术方案，其实现原理和技术效果类似，本实施例此处不再赘述。

35 参考图 8，其示出了适于用来实现本公开实施例的电子设备 700 的结构示意图，该电子设备 700 可以为终端设备或服务器。其中，终端设备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、个人数字助理(Personal Digital Assistant, 简称 PDA)、平板电脑(Portable Android Device, 简称 PAD)、便携式多媒体播放器(Portable Media Player, 简称 PMP)、车

载终端（例如车载导航终端）等等的移动终端以及诸如数字 TV、台式计算机等等的固定终端。图 8 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 8 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（Read Only Memory，简称 ROM）702 中的程序或者从存储装置 708 加载到随机访问存储器（Random Access Memory，简称 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（Liquid Crystal Display，简称 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。虽然图 8 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一——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。

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的。

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，使得该电子设备执行上述实施例所示的方法。

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（Local Area Network，简称 LAN）或广域网（Wide Area Network，简称 WAN）一连接到用户计算机，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，例如，第一获取单元还可以被描述为“获取至少两个网际协议地址的单元”。

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可以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。

第一方面，根据本公开的一个或多个实施例，提供了一种视频处理方法，包括：
获取初始视频片段；

将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率；其中，所述机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；

根据所述变速倍率对所述初始视频片段进行变速，得到目标视频片段。

根据本公开的一个或多个实施例，所述获取初始视频片段，包括：

获取初始视频；

通过对所述初始视频进行分镜，得到所述初始视频片段，其中，所述初始视频片段为所述初始视频中的分镜片段。

根据本公开的一个或多个实施例，在得到目标视频片段之后，还包括：

5 通过在所述初始视频中将所述初始视频片段替换为所述目标视频片段，得到目标视频。

根据本公开的一个或多个实施例，所述通过对所述初始视频进行分镜，得到所述初始视频片段，包括：

对所述初始视频进行分镜，得到所述初始视频中的分镜片段；

将所述分镜片段分割为多个视频片段，将其中任一视频片段确定为所述初始视频片段。

10 根据本公开的一个或多个实施例，所述将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率，包括：

提取所述初始视频片段的视频帧；

将所述初始视频片段的视频帧输入机器学习模型中，获取所述初始视频片段对应的评价参数，所述评价参数为用于衡量所述初始视频片段适宜变速程度的大小的参数；

15 根据所述初始视频片段对应的评价参数、以及预设评价参数与变速倍率对照关系，确定所述初始视频片段的变速倍率。

根据本公开的一个或多个实施例，在所述获取所述初始视频片段对应的评价参数后，还包括：

20 对所述初始视频片段对应的评价参数以及其相邻的其他初始视频片段对应的评价参数进行平滑处理，将平滑处理后的各评价参数确定为对应的各初始视频片段最终的评价参数；或者

在所述确定所述初始视频片段的变速倍率后，还包括：

对所述初始视频片段的变速倍率以及其相邻的其他初始视频片段的变速倍率进行平滑处理，将平滑处理后的各变速倍率确定为对应的各初始视频片段最终的变速倍率。

25 根据本公开的一个或多个实施例，在所述确定所述初始视频片段的变速倍率后，还包括：获取所述初始视频片段中的初始语速；

根据所述初始语速以及所述初始视频片段的变速倍率，确定所述初始视频片段按照所述变速倍率播放时的目标语速；

若所述目标语速超过预设语速阈值，则根据所述预设语速阈值调整所述变速倍率。

30 第二方面，根据本公开的一个或多个实施例，提供了一种模型训练方法，所述方法包括：获取多个训练数据；所述训练数据包括已标注了变速倍率的样本视频片段；

根据所述训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

根据本公开的一个或多个实施例，所述获取多个训练数据，包括：

35 获取已标注了变速倍率的样本视频片段，所述样本视频片段为已被慢放或快放、且被标注有慢放标签或快放标签的视频素材，和/或，未被慢放或快放、且其各视频片段分别被标注有标识适宜变速程度的的标签的视频素材；

对所述样本视频片段提取视频帧，将所述视频帧确定为所述训练数据。

第三方面，根据本公开的一个或多个实施例，提供了一种视频处理设备，包括：

获取单元，用于获取初始视频片段；

处理单元，用于将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率；其中，所述机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；

变速单元，用于根据所述变速倍率对所述初始视频片段进行变速，得到目标视频片段。

- 5 根据本公开的一个或多个实施例，所述获取单元在获取初始视频片段时，用于：
获取初始视频；

通过对所述初始视频进行分镜，得到所述初始视频片段，其中，所述初始视频片段为所述初始视频中的分镜片段。

根据本公开的一个或多个实施例，所述变速单元在得到目标视频片段之后，还用于：

- 10 通过在所述初始视频中将所述初始视频片段替换为所述目标视频片段，得到目标视频。

根据本公开的一个或多个实施例，所述获取单元在通过对所述初始视频进行分镜，得到所述初始视频片段时，用于：

对所述初始视频进行分镜，得到所述初始视频中的分镜片段；

将所述分镜片段分割为多个视频片段，将其中任一视频片段确定为所述初始视频片段。

- 15 根据本公开的一个或多个实施例，所述处理单元在将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率时，用于：

提取所述初始视频片段的视频帧；

将所述初始视频片段的视频帧输入机器学习模型中，获取所述初始视频片段对应的评价参数，所述评价参数为用于衡量所述初始视频片段适宜变速程度的大小的参数；

- 20 根据所述初始视频片段对应的评价参数、以及预设评价参数与变速倍率对照关系，确定所述初始视频片段的变速倍率。

根据本公开的一个或多个实施例，所述处理单元在获取所述初始视频片段对应的评价参数后，还用于：

- 25 对所述初始视频片段对应的评价参数以及其相邻的其他初始视频片段对应的评价参数进行平滑处理，将平滑处理后的各评价参数确定为对应的各初始视频片段最终的评价参数；或者

在所述确定所述初始视频片段的变速倍率后，还包括：

对所述初始视频片段的变速倍率以及其相邻的其他初始视频片段的变速倍率进行平滑处理，将平滑处理后的各变速倍率确定为对应的各初始视频片段最终的变速倍率。

- 30 根据本公开的一个或多个实施例，所述处理单元在确定所述初始视频片段的变速倍率后，还用于：

获取所述初始视频片段中的初始语速；

根据所述初始语速以及所述初始视频片段的变速倍率，确定所述初始视频片段按照所述变速倍率播放时的目标语速；

- 35 若所述目标语速超过预设语速阈值，则根据所述预设语速阈值调整所述变速倍率。

第四方面，根据本公开的一个或多个实施例，提供了一种模型训练设备，所述设备包括：获取单元，用于获取多个训练数据；所述训练数据包括已标注了变速倍率的样本视频片段

训练单元，用于根据所述训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

根据本公开的一个或多个实施例，所述获取单元在获取单元在获取多个训练数据时，用于：

- 5 获取已标注了变速倍率的样本视频片段，所述样本视频片段为已被慢放或快放、且被标注有慢放标签或快放标签的视频素材，和/或，未被慢放或快放、且其各视频片段分别被标注有标识适宜变速程度的标签的视频素材；

对所述样本视频片段提取视频帧，将所述视频帧确定为所述训练数据。

- 10 第五方面，根据本公开的一个或多个实施例，提供了一种电子设备，包括：至少一个处理器和存储器；

所述存储器存储计算机执行指令；

所述至少一个处理器执行所述存储器存储的计算机执行指令，使得所述至少一个处理器执行如上第一方面、第二方面以及第一方面、第二方面各种可能的设计所述的视频处理方法。

- 15 第六方面，根据本公开的一个或多个实施例，提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机执行指令，当处理器执行所述计算机执行指令时，实现如上第一方面、第二方面以及第一方面、第二方面各种可能的设计所述的视频处理方法。

第七方面，根据本公开的一个或多个实施例，提供一种计算机程序产品，包括计算机指令，该计算机指令被处理器执行时实现如上第一方面、第二方面以及第一方面、第二方面各种可能的设计所述的视频处理方法。

- 20 第八方面，根据本公开的一个或多个实施例，提供了一种计算机程序，当处理器执行计算机程序时，实现如上第一方面、第二方面以及第一方面、第二方面各种可能的设计所述的视频处理方法。

- 25 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。

- 30 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。

- 35 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，其特征在于，包括：

获取初始视频片段；

5 将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率；其中，所述机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；

根据所述变速倍率对所述初始视频片段进行变速，得到目标视频片段。

2、根据权利要求1所述的方法，其特征在于，所述获取初始视频片段，包括：

获取初始视频；

10 通过对所述初始视频进行分镜，得到所述初始视频片段，其中，所述初始视频片段为所述初始视频中的分镜片段。

3、根据权利要求2所述的方法，其特征在于，在所述得到目标视频片段之后，还包括：通过在所述初始视频中将所述初始视频片段替换为所述目标视频片段，得到目标视频。

15 4、根据权利要求2所述的方法，其特征在于，所述通过对所述初始视频进行分镜，得到所述初始视频片段，包括：

对所述初始视频进行分镜，得到所述初始视频中的分镜片段；

将所述分镜片段分割为多个视频片段，将其中任一视频片段确定为所述初始视频片段。

5、根据权利要求1所述的方法，其特征在于，所述将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率，包括：

20 提取所述初始视频片段的视频帧；

将所述初始视频片段的视频帧输入所述机器学习模型中，获取所述初始视频片段对应的评价参数，所述评价参数为用于衡量所述初始视频片段适宜变速程度的大小的参数；

根据所述初始视频片段对应的评价参数、以及预设评价参数与变速倍率对照关系，确定所述初始视频片段的变速倍率。

25 6、根据权利要求5所述的方法，其特征在于，在所述获取所述初始视频片段对应的评价参数后，还包括：

对所述初始视频片段对应的评价参数以及其相邻的其他初始视频片段对应的评价参数进行平滑处理，将平滑处理后的各评价参数确定为对应的各初始视频片段最终的评价参数；或者

30 在所述确定所述初始视频片段的变速倍率后，还包括：

对所述初始视频片段的变速倍率以及其相邻的其他初始视频片段的变速倍率进行平滑处理，将平滑处理后的各变速倍率确定为对应的各初始视频片段最终的变速倍率。

7、根据权利要求5或6所述的方法，其特征在于，在所述确定所述初始视频片段的变速倍率后，还包括：

35 获取所述初始视频片段中的初始语速；

根据所述初始语速以及所述初始视频片段的变速倍率，确定所述初始视频片段按照所述变速倍率播放时的目标语速；

若所述目标语速超过预设语速阈值，则根据所述预设语速阈值调整所述变速倍率。

8、一种模型训练方法，其特征在于，包括：

获取多个训练数据；所述训练数据包括已标注了变速倍率的样本视频片段；

根据所述训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

9、根据权利要求8所述的方法，其特征在于，所述获取多个训练数据，包括：

- 5 获取所述已标注了变速倍率的样本视频片段，所述样本视频片段为已被慢放或快放、且被标注有慢放标签或快放标签的视频素材，和/或，未被慢放或快放、且其各视频片段分别被标注有标识适宜变速程度的的标签的视频素材；

对所述样本视频片段提取视频帧，将所述视频帧确定为所述训练数据。

10、一种视频处理设备，其特征在于，包括：

获取单元，用于获取初始视频片段；

- 10 处理单元，用于将所述初始视频片段输入机器学习模型中，并根据所述机器学习模型的输出结果确定所述初始视频片段的变速倍率；其中，所述机器学习模型已基于已标注了变速倍率的样本视频片段进行了训练；

变速单元，用于根据所述变速倍率对所述初始视频片段进行变速，得到目标视频片段。

11、一种模型训练设备，其特征在于，包括：

- 15 获取单元，用于获取多个训练数据；所述训练数据包括已标注了变速倍率的样本视频片段；

训练单元，用于根据所述训练数据对机器学习模型的初始模型进行训练，得到训练后的机器学习模型。

12、一种电子设备，其特征在于，包括：至少一个处理器和存储器；

- 20 所述存储器存储计算机执行指令；

所述至少一个处理器执行所述存储器存储的计算机执行指令，使得所述至少一个处理器执行如权利要求1-9任一项所述的方法。

13、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中存储有计算机执行指令，当处理器执行所述计算机执行指令时，实现如权利要求1-9任一项所述的方法。

- 25 14、一种计算机程序产品，包括计算机指令，其特征在于，该计算机指令被处理器执行时实现如权利要求1-9任一项所述的方法。

15、一种计算机程序，其特征在于，所述计算机程序被处理器执行时实现如权利要求1-9任一项所述的方法。

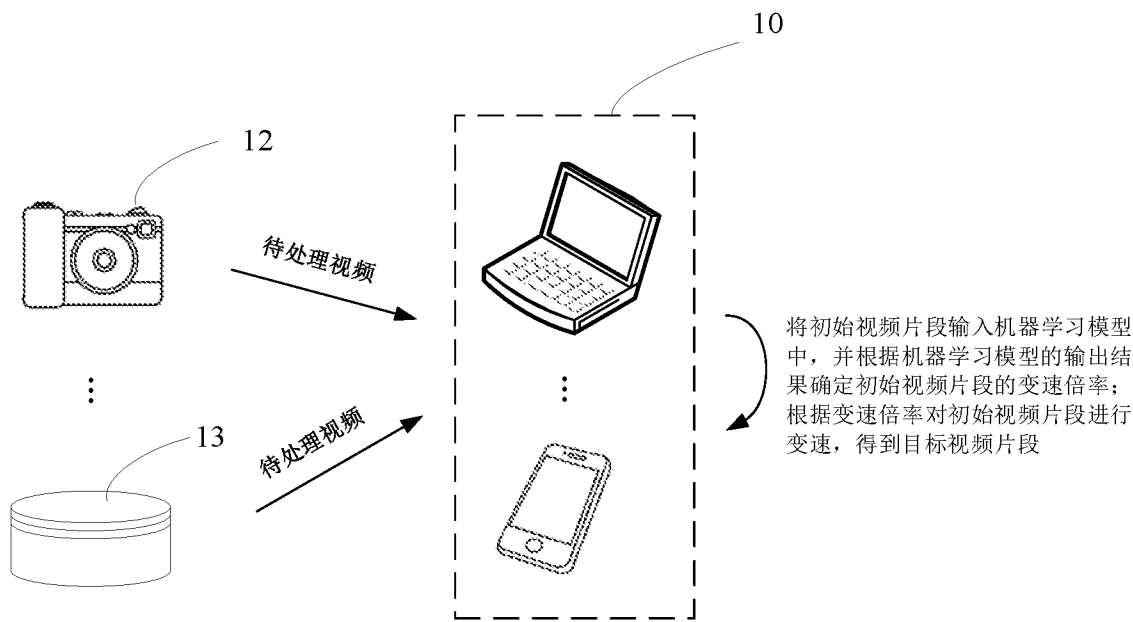


图 1

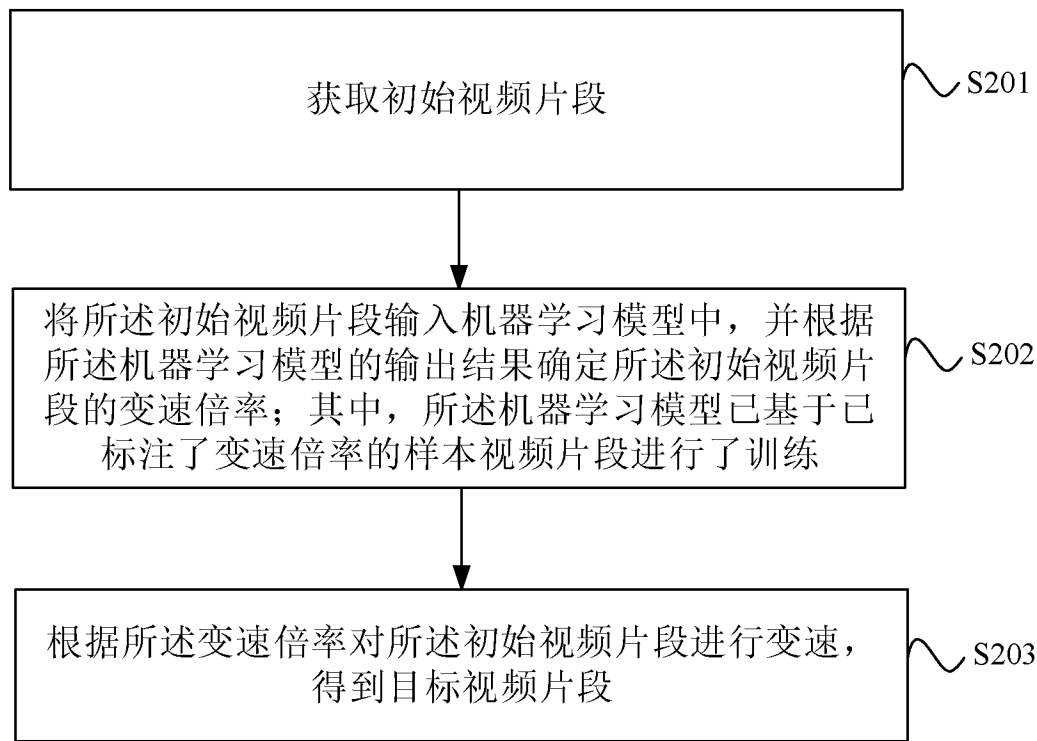


图 2

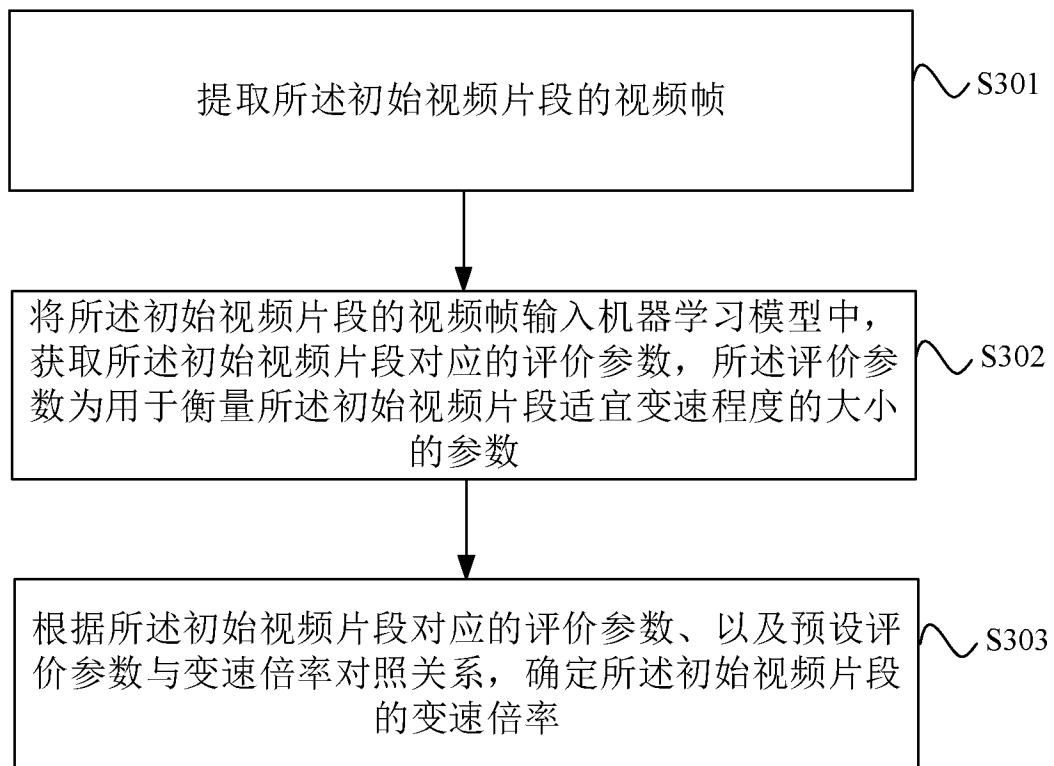


图 3

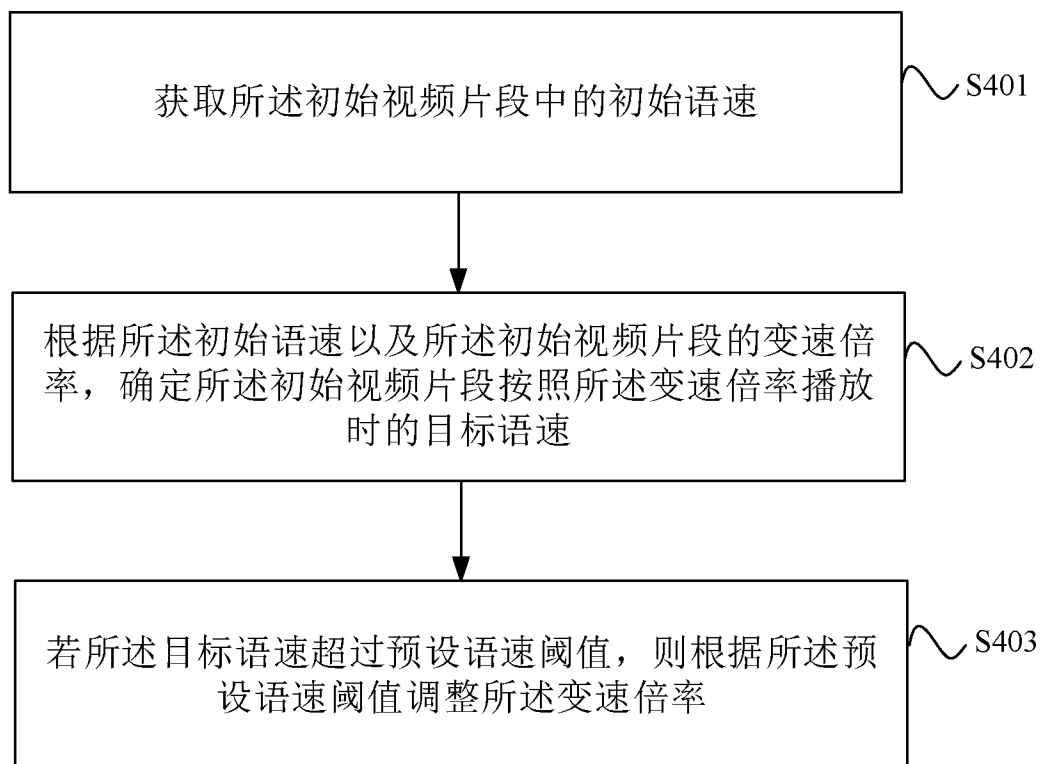


图 4

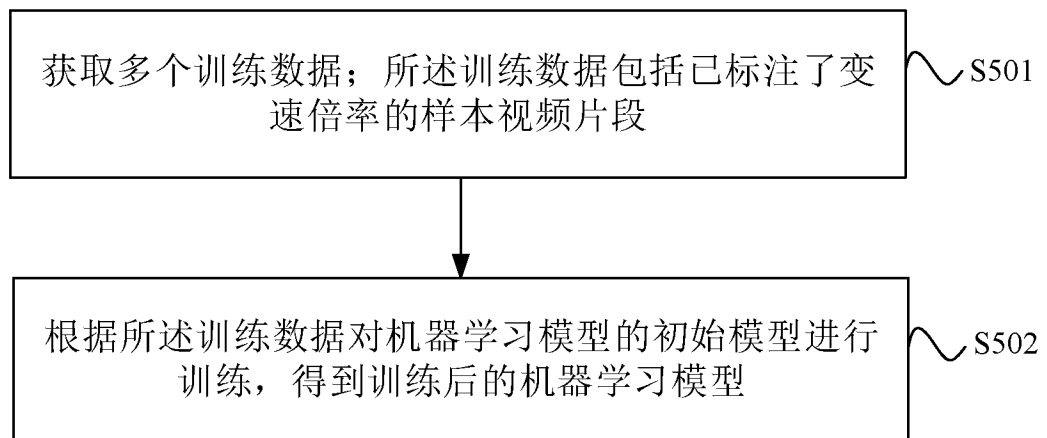


图 5

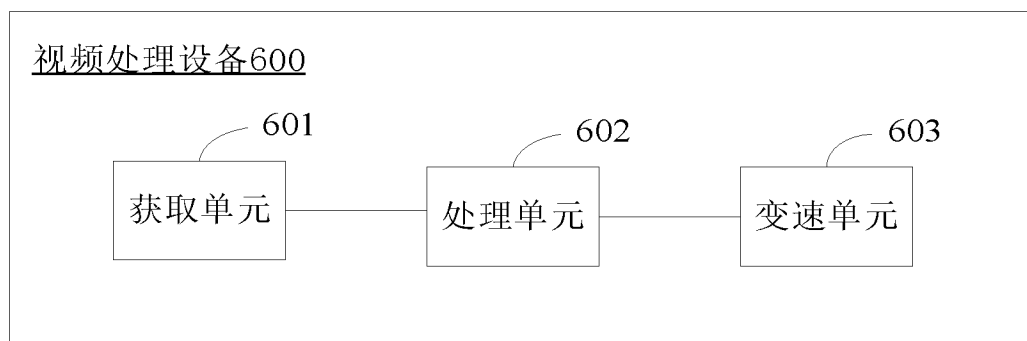


图 6

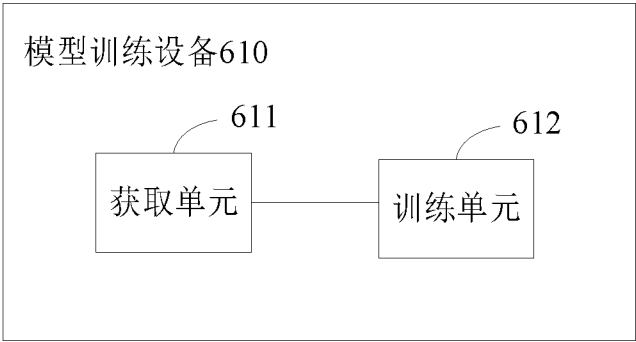


图 7

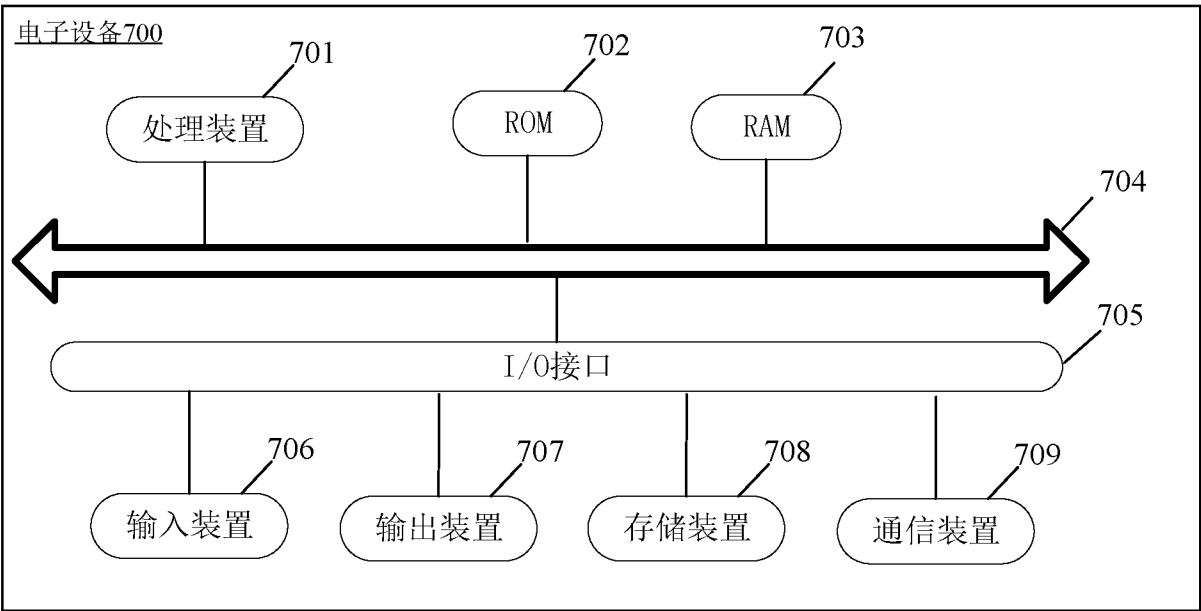


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/096646

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/44(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)

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, WPABSC, ENTXTC, WPABS, ENTXT, CNKI, IEEE: 视频, 编辑, 播放, 回放, 变速, 快进, 快放, 快退, 慢放, 慢镜头, 自动, 自适应, 训练, 模型, video, edit+, play, playback, variable 3d speed, chang+ 3d speed, forward, fast, back, slow 3d speed, slow 1w motion, auto+, adapt+, learn+, model

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 111771384 A (GOOGLE INC.) 13 October 2020 (2020-10-13) description, paragraphs 0015-0092 and 0175-0182, and figures 1A-3A	1-5, 7-15
Y	CN 111771384 A (GOOGLE INC.) 13 October 2020 (2020-10-13) description, paragraphs 0015-0092 and 0175-0182, and figures 1A-3A	6
Y	CN 110730387 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 24 January 2020 (2020-01-24) description, paragraphs 0128-0131	6
A	CN 104735385 A (BEIJING XIAOMI TECHNOLOGY CO., LTD.) 24 June 2015 (2015-06-24) entire document	1-15
A	CN 104506947 A (FUZHOU UNIVERSITY) 08 April 2015 (2015-04-08) entire document	1-15
A	US 7127127 B2 (MICROSOFT CORPORATION) 24 October 2006 (2006-10-24) 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

“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

12 July 2022

Date of mailing of the international search report

27 July 2022

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/096646

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111771384	A	13 October 2020	KR	20200124266	A	02 November 2020
				EP	3759931	A1	06 January 2021
				WO	2019168868	A1	06 September 2019
				US	2019268632	A1	29 August 2019
				JP	2021515479	W	17 June 2021
				IN	202047040965	A	25 September 2020
CN	110730387	A	24 January 2020	HK	40020144	A0	23 October 2020
CN	104735385	A	24 June 2015	US	2016293209	A1	06 October 2016
				WO	2016155228	A1	06 October 2016
				BR	112015028827	A2	25 July 2017
				JP	2017514427	A	01 June 2017
				MX	2015015112	A	05 December 2016
				KR	20160127630	A	04 November 2016
				EP	3076664	A1	05 October 2016
				RU	2015150036	A	25 May 2017
				KR	101712306	B1	03 March 2017
				JP	6374598	B2	15 August 2018
				MX	355081	B	04 April 2018
				CN	104735385	B	12 October 2018
				IN	201503686	P2	13 October 2017
				US	10217487	B2	26 February 2019
CN	104506947	A	08 April 2015	CN	104506947	B	05 September 2017
US	7127127	B2	24 October 2006	US	2004175058	A1	09 September 2004

国际检索报告

国际申请号

PCT/CN2022/096646

A. 主题的分类

H04N 21/44 (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)

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, WPABSC, ENTXTC, WPABS, ENTXT, CNKI, IEEE: 视频, 编辑, 播放, 回放, 变速, 快进, 快放, 快退, 慢放, 慢镜头, 自动, 自适应, 训练, 模型, video, edit+, play, playback, variable 3d speed, chang+ 3d speed, forward, fast, back, slow 3d speed, slow 1w motion, auto+, adapt+, learn+, model

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
X	CN 111771384 A (谷歌有限责任公司) 2020年10月13日 (2020 - 10 - 13) 说明书第0015-0092、0175-0182段, 附图1A-3A	1-5、7-15
Y	CN 111771384 A (谷歌有限责任公司) 2020年10月13日 (2020 - 10 - 13) 说明书第0015-0092、0175-0182段, 附图1A-3A	6
Y	CN 110730387 A (腾讯科技深圳有限公司) 2020年1月24日 (2020 - 01 - 24) 说明书第0128-0131段	6
A	CN 104735385 A (小米科技有限责任公司) 2015年6月24日 (2015 - 06 - 24) 全文	1-15
A	CN 104506947 A (福州大学) 2015年4月8日 (2015 - 04 - 08) 全文	1-15
A	US 7127127 B2 (MICROSOFT CORPORATION) 2006年10月24日 (2006 - 10 - 24) 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

“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

2022年7月12日

国际检索报告邮寄日期

2022年7月27日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

曹珊珊

电话号码 027-59371318

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/096646

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	111771384	A	2020年10月13日	KR	20200124266	A	2020年11月2日
				EP	3759931	A1	2021年1月6日
				WO	2019168868	A1	2019年9月6日
				US	2019268632	A1	2019年8月29日
				JP	2021515479	W	2021年6月17日
				IN	202047040965	A	2020年9月25日
CN	110730387	A	2020年1月24日	HK	40020144	A0	2020年10月23日
CN	104735385	A	2015年6月24日	US	2016293209	A1	2016年10月6日
				WO	2016155228	A1	2016年10月6日
				BR	112015028827	A2	2017年7月25日
				JP	2017514427	A	2017年6月1日
				MX	2015015112	A	2016年12月5日
				KR	20160127630	A	2016年11月4日
				EP	3076664	A1	2016年10月5日
				RU	2015150036	A	2017年5月25日
				KR	101712306	B1	2017年3月3日
				JP	6374598	B2	2018年8月15日
				MX	355081	B	2018年4月4日
				CN	104735385	B	2018年10月12日
				IN	201503686	P2	2017年10月13日
				US	10217487	B2	2019年2月26日
CN	104506947	A	2015年4月8日	CN	104506947	B	2017年9月5日
US	7127127	B2	2006年10月24日	US	2004175058	A1	2004年9月9日